

Factor and Organizational Substitutions to Minimize Costs in the Navy

Ugurhan G. Berkok¹, Jeffrey Penney² and Steven Kivinen³

Prepared by:

Ugurhan Berkok

Royal Military College of Canada and Queen's University

Project Manager: Robert M.H. Burton

Contract Number: PA12022

Contract Scientific Authority: Binyam Solomon

The scientific or technical validity of this Contract Report is entirely the responsibility of the contractor and the contents do not necessarily have the approval or endorsement of Defence R&D Canada.

Defence R&D Canada – CORA

Contract Report

DRDC-RDDC-2013-C10

December 2013

¹ Dept. of Management & Economics, Royal Military College of Canada and Dept. of Economics, Queen's University

² Dept. of Economics, Queen's University

³ Dept. of Economics, Queen's University

Principal Author

Original signed by Ugurhan Berkok [Royal Military College of Canada and Queen's University]

Ugurhan Berkok [Royal Military College of Canada and Queen's University]

Approved by

Original signed by Robert M.H. Burton

Robert M.H. Burton

Section Head, Joint Systems Analysis

Approved for release by

Original signed by Eric Fournier

Eric Fournier

Chief Scientist

Abstract

This report focuses on potential input substitutions in a navy. The first part introduces elements of economic analysis of input substitutions in production theory, incorporating such critical issues as risk, time and transaction costs as well as widening the scope of substitutions beyond the simple labour for capital framework as consistent with global practice. The second part is divided into four sections. The first two sections cover substitutions within the navy itself and its fleet followed by substitutions at a more general level within the government. The next two sections cover the ship design at a technical level and affecting procurement and eventually the operations and maintenance costs and, finally, beyond the government, outsourcing seen as organizational substitution. Each section contains examples from various navies. Although automation arises in such major cases as modularity, innovative substitutions introduced at the design stage and hitherto unconsidered outsourcing practices emerge as part of menus available to defence planners. Two major questions arise for future research. One is the nature and extent of interaction between substitutions. The other is about the menu of substitutions available to smaller navies.

Executive summary

This study focuses on potential input substitutions in a navy. Improvement in capabilities and reduction in costs, or a combination, may be the objectives of defence planners. Reducing the cost of performing a task has traditionally been realized by automation or the substitution of costly human resources for lower cost equipment. For example, mine detection in a harbour can be performed by at least three methods: Robotic detection devices, properly equipped navy divers, and well-trained and camera-equipped bottlenose dolphins. This static cost minimization does not require an all-or-nothing substitution. Rather, considering that dolphins may not succeed in all environments, hybrid teams can be used for scouring a given harbour at minimum cost. The first part of the study covers a review of economic analysis on substitutions where substitution is understood to be wider in scope, as consistent with global practice, than mere labour for capital substitutions if the objective is one of cost minimization.

Static input substitution principle that cheaper inputs should replace more expensive ones won't apply to navy platforms simply because these latter are built to last for decades. Beyond strategic reasons that may necessitate them, expensive current inputs may save lifetime costs by enabling future cost-reducing substitutions. Under rapid technological change in electronics, new materials and innovative designs, the scope of substitution is wider. Since defence planners have correctly shifted their cost focus from acquisition to lifetime costs, intertemporal substitutions enabled by new technologies, such as current modular design relaxing the limited mission range of a naval ship by enabling it to switch mission modules at short notice, may save costs by narrowing the range of ships in a fleet. This observation alone must guide fleet choices of navy planners. The study offers several examples in Part 2 of the study.

To the extent that substitutions are external to the navy, they are implemented by outsourcing, procurement and employment contracts. The transaction cost analysis then sheds light on emerging contracting issues. Whereas contract design problem precedes the transaction under contract, the transaction cost and the accompanying contract governance concerns succeed the transaction. Adaptation to new conditions such as arising from contract incompleteness, both exogenously and endogenously, require proper governance structures for cost minimization purposes. These potential contract design and implementation costs suggest caution in implementing substitutions that could otherwise have been reducing lifetime costs.

The second part of the study categorizes many potential navy input substitutions into four sections.

Firstly, we consider substitutions exclusively related to navy ships. The simplest example of the substitution of personnel for equipment is crew rotation, which consists of equipment-personnel substitution with a significant influence on lifetime ship costs. Furthermore, a fleet's structure in terms of vessel combinations, beyond that required by strategic considerations, allows substitution across ships to minimize fleet cost. It must be noted, though, this substitution is more suitable in the case of large navies with a variety of ships such as in the case of US Navy. Finally, advances in robotics technologies widen the scope of automation and task redesign to incorporate such technologies paves the way to automation, especially within modular overall designs. Whereas an adjustment of fleet

structure with a narrow array of ships may not be an option for Canadian Navy, crew rotation is already being implemented and further automation, at the level of the new German frigates that reduced personnel by about 20%, should also be feasible for the upcoming Canadian Surface Combatants, Arctic Offshore Patrol Ships and support ships.

Secondly, scope economies can be exploited where government departments can directly cooperate or simply coordinate to reduce overall costs. Areas within naval defence that can be commissioned to other departments include supply chain management and logistics. In Canada, this latter is a largely outsourced function where a transaction cost analysis may yield in favour of substituting such outsourcing for a government agency. Similarly, forward bases, hubs or stations can conceivably be run by a purchasing and servicing department of the government. Even such a directly operational area as search and rescue can be run by the coastguard or the navy instead of the air force. All substitutions in this second group are feasible in Canada, pending a cost-benefit analysis.

Thirdly, since a naval ship is a complex system of systems including navigation, propulsion, combat, hotel, communications, in-service support and any system corresponding to further tasks such as transport, hospital services and command and control, the personnel for equipment substitutions can be conceptualized across the subsystems or modules as well as within them. Modularity, incorporated at the design stage, may not only increase timely switches between capabilities but may also affect procurement, maintenance and operating costs. For instance, mission modules stored at forward hubs or onboard, while increasing ship costs may reduce the number of trips to the home port significantly. Moreover, since modularity severs the jointness of various capabilities onboard the platform, a given module can be upgraded relatively independently of others. The new Canadian Surface Combatants and Arctic Offshore Patrol Ships can be designed to carry various mission modules whereas the support ships can be similarly designed as modular.

Fourthly, outsourcing certain functions may reduce ownership costs, either by implementing new technologies or by decreasing X-inefficiency that conceivably exists within a military organization more preoccupied with operational effectiveness than efficiency. This section considers the relationship between deployment and substitutions. Continuous deployments, relevant for countries like the United States, pose different problems, especially in logistics with forward bases and stations, personnel rotations and fleet structures. Furthermore, several examples of outsourcing in navies of Australia, France, the U.K. and the U.S. are considered and they are drawn specifically from operational areas of naval defence rather than outsourcing in procurement and logistics.

Two major questions arise for future research. One question is the nature and extent of interaction between different substitutions. For example, how would modularity and automation by ship design interact? What can we learn from Trans-Tasman defence cooperation where New Zealand outsourced sea patrols to Australian Navy? Will this outsourcing of an operational capability generate scale diseconomies for New Zealand or force hollowing out? The other question is whether larger navies may have a wider spectrum of possibilities? A larger navy may make use of a larger number of smaller or simpler vessels instead of bigger or more complicated ones but economizing nevertheless from procurement cost savings due to longer production runs. Yet another example is the organizational

substitution with U.S. Navy's surface warfare officer retention auctions where the existence of a large group of officers and a large number of billets make it possible to run combinatorial auctions to match officers with postings while saving overall costs.

Contents

Abstract / ii

Executive summary / iii

Introduction / 1

1. Substitutions in production / 3

1.1 Production theory / 3

1.2 Transaction Costs / 5

1.3 Optimal Contracts / 6

1.4 Allocation of Inputs / 7

1.5 Time and Uncertainty / 8

2. Substitutions in practice / 9

2.1 Substitutions within navy: Factor and organizational substitutions / 9

- Easier in Navy than Army? / 9
- Cost implications of crew rotation / 10
- Fleet structure and costs / 12
- Task redesign and automation / 13

2.2 Substitutions within government / 15

- Devolution of civilian employment, recruitment and acquisition functions to other Federal departments / 15
- Supply chain for operational support and forward hubs / 15

2.3 Substitutions by ship design / 16

- Design and development cost / 16
- Modular adaptable ships / 17
- Procurement cost / 19
- Personnel and maintenance costs / 20
- Fuel costs / 20

2.4 Substitutions by outsourcing aspects of military operations / 21

- Legal considerations / 21
- Market imperfections / 24
- A simple model of outsourcing / 27
- Choosing tasks to outsource / 28
- Navy outsourcing examples / 28
- The military services industry in the USA / 30

Conclusions and extensions / 33

References / 32

This page intentionally left blank.

Introduction

Modernization in defence forces is continuously in agenda for purposes of improving capabilities or minimizing costs as allowed by technological advances. Militaries modernize by retrofitting and replacing platforms, or by modifying the composition and/or configuration of capabilities. In both cases, technological advances impose factor substitutions in light of the fact that personnel are arguably the most expensive factor over the lifetime of platforms.

Some nations are doing so rapidly: for example, China has been spending considerable resources to avail their armed forces of the latest technologies, with Japan and India having similar aspirations. For example, "... the United States has no choice but to pursue a two-pronged policy towards China. One prong is to engage China and encourage it to become a "responsible stakeholder" in the international community. The second is to engage in "prudent hedging" against competitive or aggressive behavior by China, pursuing continued engagement rather than treating the country as an enemy. ... There are many costs associated with hedging against China. In terms of military investment, this mission requires ultra-modern aerospace and naval capabilities. It is the main budgetary rationale for an advanced fighter aircraft, a new strategic bomber, new aircraft carriers and other surface combatants, stealthy unmanned aerial systems with long range and dwell time, nuclear attack submarines, and a host of C4ISR assets." (Carter & Bulkeley(2007)) Thus, modernization is a strategic imperative rather than an optional path.

The decision to modernize comes perennially with two questions: the first, what kind of military hardware to keep and which to retire due to obsolescence; the second, whether it is effective and feasible to replace manned positions in various war machines with technological substitutes.

This report focuses on the latter question as well as related issues and will guide towards solutions using economic analysis. In the first part, the literature review covers relevant topics in the economic theory of production, contracts and outsourcing, matching, and dynamic decision-making, this latter being especially relevant for an analysis of military platforms with long service lives and perpetually competing against shocks from progressing technologies. The theme is restricted to cost minimization through "substitution" and thus leaving a full cost-benefit analysis to military planners. Below in Part 1 we review the relevant elements of economic analysis. Production theory concentrates on cost-minimizing input bundles whereas transaction cost theory examines the cost-minimizing combination of in-house and outsourced inputs for organizing the firm's activity. Contract theory investigates the best contracts when outsourcing as well as securing various inputs and selling the products. Matching theory looks at the allocation of inputs across the organization. Finally, decision making across time involves tradeoffs over time and, especially in the case of assets with long service lives such as military platforms, determining factors such as discounting and risk have to be incorporated into economic analysis.

Naval ships exhibit long service lives in the range of several decades. Since new technologies can only be partially integrated during a traditional ship's life cycle, complex platforms and hence their input compositions can significantly be modified only at mid-life refits. Major redesigns with new technologies can only be implemented at new vintages. However, current design choices enlarge options to include short lives in the range of 10 to 15 years versus long lives with built in modularity so that fast

progressing mission and communication systems can be made available to military planners. (O'Rourke (2013)) Traditionally, though, if a personnel-for-equipment or any other substitution is required to minimize life-cycle ownership costs, such substitution may take various distinct forms. We classify them into four groups.

Firstly, substitutions within navy are nearly exclusively related to navy ships, save forward bases and stations. A brief discussion of whether substitutions are easier in the army than in the navy concentrates on the initial state of the army as boots on the ground whereas the navy being inevitably ship-centric. Crew rotation applied to selective ships in various navies is a prime example of equipment personnel substitution as well as of the impact of other factors on the decision in terms of cost implications. At a higher level of decisions, fleet structure also poses decision challenges in terms of vessel combinations required by strategic combinations. However, given that there exist various combinations to achieve the strategic objectives, the cost-minimization objective forces substitutions. Mostly at the ship level, automation is easier when tasks to be automated are simple. Although advances in robotics technologies widen the scope of automation, task redesign paves the way to automation.

Secondly, an area where scope economies could yet be exploited is where government departments can directly cooperate or simply coordinate to reduce overall costs. Apart from obvious cost savings in personnel management and procurement, a few areas within naval defence are considered potential candidates for commissioning to other government departments. Supply chain management and logistics can be devolved to or run cooperatively with a specialized government department. Even such a directly operational area as search and rescue can be run by the coastguard or a ministry in charge coastal natural resources.

Thirdly, in the design stage where latest technology can be incorporated taking into account the tasks the ship has to perform and the complexity of systems corresponding to those tasks. A naval ship is a system of systems including navigation, propulsion, combat, hotel, communications, in-service support and any system corresponding to further tasks such as transport, hospital services and command and control. Therefore, the personnel for equipment substitution can be conceptualized in terms of these subsystems or modules as well as within them. This study includes examples where a new vintage may reduce ownership costs by embodying new technologies.

Fourthly, outsourcing any of the functions may reduce ownership costs. Outsourcing may reduce costs either by implementing new technologies or by decreasing X-inefficiency that might exist within a military organization more preoccupied with operational effectiveness than efficiency. A discussion of risks, costs and benefits of outsourcing is included. Outsourcing with accompanying personnel substitutions can be instrumental for lower ownership costs. The study also includes an examination of the relationship between deployment and substitutions. For instance, especially during a surge when reserves are mobilized, civilians may temporarily step in to perform functions performed by regulars and reserves but, in support functions, such substitutions are commonplace. Continuous deployments, relevant for countries like the United States, pose different problems, especially in logistics with forward bases and stations, personnel rotations and fleet structures.

We then turn to a discussion of ancillary issues. The legal limits of privatization of various sectors of defence are explored. Whereas international agreements Canada is a signatory to are complex in nature, we simplify the analysis to the borderline case where Canada might seek the use of private military companies (herein PMCs) to conduct frontline combat operations. While there exist multiple considerations and legal opinions vary, there does not appear to be any legal limitation preventing Canada from outsourcing to PMCs. Rather, other legal considerations, such as the potential liability for the behaviour of PMC employees, prove of a higher order concern (Singer (2003)). The uncertainties of war and conflict make certain types of contracts especially incomplete, potentially granting bargaining power to subcontracted PMCs. Although reputation may serve to assuage some of these concerns, the potentially small size of the market (depending on the desired service) and the fact that performance is not always comparable between competitors diminish the impact of the reputational consideration somewhat. A simple game-theoretic model is built to illustrate the decision to outsource military procurement as an example. Outsourcing aspects of military procurement brings with it a number of different concerns. Primarily, the principal-agent problem due to imperfect monitoring of performance and to potential difficulties with contract enforcement, affect whether to outsource the task to the private sector. We conclude the section by listing various factors which tend to increase the benefits of outsourcing.

The study is organized as follows. Section 1 briefly reviews elements of economic analysis relevant to this study. Section 2 covers details of substitutions that may induce cost savings and practical examples within the context of navy ships, fleet structure and intergovernmental substitutions. Four broad sections cover, respectively, substitutions within the navy and within the government, substitution by outsourcing and lastly ship and fleet designs. Finally, the Conclusions and Extensions section points out to relevant directions to deepen the study of examples in Section 2.

1. Substitutions in production

This review focuses on several aspects of efficiency in economic analysis: production costs, transaction costs, contract design, intertemporal allocation of resources in the presence of risk, and intertemporal substitution. Strategies for minimizing costs are also suggested.

1.1 Production Theory

The basic idea underlying production theory is that such economic agents as firms and governments can combine inputs in various ways to produce goods and services. These productive agents purchase inputs in factor markets and sell the outputs in goods and services markets. For instance, the Royal Canadian Navy (RCN) purchases inputs, such as labour, capital and intermediate inputs such as fuel from the usual inputs markets. Some inputs, such as skills and human capital, are often produced within the organization. But differing from regular goods and services, the main output of RCN, i.e. national security, is not sold on the market but consumed directly by all Canadians whereas funds to purchase

and produce the inputs flow from taxes while the basic process of production remaining identical to the production of goods and services in general.

Several questions arise when thinking about productive efficiency. Is the correct amount of output being produced? Is the correct mix of inputs being used? While advanced treatments of production theory can be found in microeconomics textbooks (e.g. Jehle & Reny (2011), Mas-Colell, Winston & Green (1995)), the ideas can be clearly and simply explained using the concept of production function, which describes the relationship between the quantities of inputs and outputs. One must underline the scope of the concept of output. Often and especially in defence, a range of outputs are produced jointly. For instance, RCN would produce deterrence jointly while intercepting illegal immigrant or fishing boats as well as projecting power that would assist government's international policy. Hildebrandt (1999) and Rohlfs (2005) provide thorough examples of estimating military production functions, the former by an air force and the latter by an army. Applying the concept of defence output developed in Hildebrandt, if the RCN wishes to know the impact of various strategies on illegal fishing in Canadian waters, one can analyze data on the incidence of such activity against data on the number of patrol vessels, the use of radar, the use of surveillance and reconnaissance aircraft as partially determining the national security output in question. Thus, one can relate the level of illegal fishing to naval inputs and recover a production function.

Different inputs have different effects on production. For example, 2 soldiers with 2 guns each, or 4 soldiers sharing 2 guns are likely to be much less effective than 3 soldiers with one gun each. On the other hand, a soldier using a Type A rifle may be similarly as productive as a soldier using a Type B rifle. The idea expressed by the example is the substitutability of inputs. Whereas similar guns are substitutes, soldiers and guns have typically been complements but, with advancements in technology, soldiers are being progressively replaced by guns. One might easily invoke the recent replacement of piloted aircraft with drones as an example.

There are several measures of input substitutability in production processes. Jehle & Reny (2011) provides precise definitions of such measures. One important measure is the marginal rate of technical substitution. An intuitive explanation stems from what happens when input quantities are varied. If, resuming the simple example above, one were to reduce the number of guns by one, how many soldiers do I need to achieve the same amount of security? There would probably be no drop in security if a gun were to be taken away from the team of 2 soldiers with 2 guns each as the two soldiers would still be equipped with sufficient weaponry whereas there would be a drastic fall of output if a gun were to be taken away from the team of 4 soldiers sharing 2 guns or the one where 3 soldiers with one gun each. In the same vein of argument, if Type A guns were to be reduced by one, how many Type B guns would return the same level of security as before? The answer would certainly differ from the case where various guns exhibit different capabilities. Of course, in all these examples, there are cost implications in the sense that input combinations will project different overall costs. The rest of this report applies these main ideas to a naval force.

Given a production function constrained by the available technologies of the day, the current input and output prices, and the prevailing objectives of the navy one can derive an optimal set of inputs and expected outputs. The simpler problem asks: Given a production function, the input prices and the expected output, are inputs chosen so as to minimize costs? In fact, this is the quintessential efficiency question as governments trade off the expected output versus costs or, in other words, defence objectives versus fiscal realities.

The generic cost minimization problem is well understood. If additional security is costlier than what would be acceptable fiscally, the prevailing level of security provision is then optimal. Technically speaking, inputs should be chosen such that marginal cost of security should equal its marginal benefit. However, the costing of this level of security presumes that amongst available combinations of inputs, all assuring the same level of security, the least costly combination is chosen. Thus, there is an underlying cost minimization problem in resource allocation that, whatever the level of provision, it should be cost effective. This cost minimization problem is solved if inputs are chosen in such a way that the ratio of any two input prices equals the marginal rate of technical substitution between those inputs. Intuitively speaking, a dollar spent in a particular military capability should produce the same additional security as in any other capability because, if not, the total security would increase if dollars are transferred to more productive uses. For example, US nuclear deterrent is a three-pronged capability with navy, air force or army launched nuclear warheads. If, for instance, one delivery mechanism becomes more affordable, one would expect a substitution of away from the more to less costly mechanism and hence a reduction of the total cost of the overall deterrent. Thus, changes in input prices will affect the optimal input bundle when inputs are substitutes. Of course, cutting costs can always be realized by lowering security. Hartley (2004) discusses this issue with regard to the U.K. military.

1.2 Transaction Costs

Should the RCN privatize its management of ship procurement? When answering such a relevant question, especially in the light of the British seriously considering such a privatization (Chuter (2013), Defence Management (2013)), two basic issues arise.

First, one might wonder whether a private firm is more productively efficient. If so, it is tempting to vigorously consider privatization. Since the output is the acquisition of naval vessels as needed by the navy, privatization is input substitution or, to be exact, organizational substitution. However, Coase (1937) cautions against hasty action: One must consider the relative costs of using the market system versus the cost of using an intra-firm organization. Franck & Mélése (2008) provide a detailed introduction to transaction costs and military production. To see why using the market may be costlier than producing the same service in house, even with low production efficiency, consider the following example.

Suppose there are two ways to proceed with naval vessels acquisition, each requiring different skill sets. It's worth noting that the expected outcome is identical whereas competing procurement organizations differ. For instance, the government procurement agency may use a technology consistent with the hierarchical structure of government and the privatized agency another one consistent with a technology consistent with decentralized markets. This latter one would require high-powered incentives to motivate economic agents whereas the former would motivate them by career path incentives. If the private alternative is chosen, the contract will normally stipulate incentives based on cost and time overruns. However, due to cost overruns from a source not specified in the contract, contract implementation issues may arise ex post. For instance, the navy may acutely need the platform and, having irreversibly committed to the acquisition, accept the higher price without being able to recover any cost overruns from the agency due to contract incompleteness.

Three concepts are inherent to transaction costs hypothesis. First, there is an investment in human capital specific to the transaction. This is known as asset specificity. Second, the contract does not have

a contingency for every event that may occur. Although this contract incompleteness may not necessarily be exogenous, it does exist in almost all contracts (Halonen-Akatwijuka & Hart (2013)). Finally, the private firm may act opportunistically in the sense of passing on all cost overruns. Even though the private firm can exhibit productive efficiency, i.e. transforming inputs into outputs, the contractual relationship leaves room for the firm to exploit the purchaser. This well-known hold-up problem is especially relevant in defence related markets that happen to be thin, i.e. with few buyers and few sellers.

Transaction costs matter critically when deciding whether to use the market. There are many transaction costs besides asset specificity (Franck & Mélèse (2008)). Three important sources of transaction costs are search and information costs, bargaining costs, and property rights enforcement costs. Beyond the privatization of procurement management, the procurement agency itself may have difficulty assessing which firms are the most productively efficient. Auctions may help overcome this hurdle, though the time and resources required to setting up an auction generate transaction costs. It is worth noting that these latter transaction costs arise under all procurement organizational form, whether public or private. Williamson (1999) discusses the functioning of bureaucracies from a transaction cost perspective.

Finally, when logistics or procurement services are outsourced, if the supply to the military authority in operational conditions arrives late or not at all, the military authority will have no problem determining that the contractor is in breach of its obligations under the contract. The result is that the appropriate consequences to the contractor, as stipulated in the contract, will be applied. But what if what's involved is the physical protection of diplomats or a military authority's business sites? Here it may be much more difficult for the military authority to assess compliance with the requirements of the contract. It is extremely important then to consider, before the contract is signed, the military authority's performance expectations, so that non-performance can adequately be assessed, with the appropriate consequences clearly articulated in the contract.

1.3 Optimal Contracts

Asset specificity is a problem when contracts are incomplete. However, even if incompleteness may not be a problem, other issues remain with contracting. In the literature on agency contracting, the principal would have difficulty conditioning contracts on some unobservable variables such as some actions of the agent or on private information held by the agent. However, if the agent's actions are known to be correlated with observables (such as completion time, etc.) then the principal can design a contract contingent on those observables.

For example, in a military aircraft acquisition with fixed price contracts, the contractor would like to reduce costs while executing the contract. If such cost cutting measures lead to less planes being built or lengthy completion times then it is optimal to condition the price on total number of planes built or the date they are received. In reality, such cost and time overrun incentives are standard practice. The idea expressed in the example is the implementability or incentive compatibility of the contract beyond the efficiency in a naïve perspective. The best contract is the most efficient of all incentive compatible contracts. Mas-Colell, Winston & Green (1995) provides an advanced treatment of these issues.

1.4 Allocation of Inputs

Most countries with navies have many locations for their naval bases or postings. This observation invokes spatial distribution considerations in the allocation of personnel. Which sailors should be allocated to which posts? Here we review the methods of matching when (i) the post has no preference, but the sailor does, and (ii) both the sailor and the posts have preferences over each other. When the post does not care which sailors it receives, the RCN can simply auction off assignments. For example, in the case of RCN, a sailor from Halifax may prefer to be stationed in Halifax instead of Victoria. If the RCN is indifferent to the location, the sailor may be permitted to pay the RCN to move. Is this of mutual benefit? Personnel will then cost less to RCN, but is the sailor better off? So long as the sailor does not over-bid then s/he is also better off. This outcome critically depends on designing the auction in question. A major finding of auction theory is that a second-price sealed bid auction has bidders bidding their true evaluations. Vickrey (1961) describes the outcome of such an auction. Strategically, bidders behave similarly to an English (or ascending bid) auction where bidders orally and openly disclose their bids. In a Vickrey auction, bidding one's true value is the best s/he can do. Another finding discovered by Vickrey (1961) is the revenue equivalence theorem. Essentially, the theorem says that all auctions (that satisfy basic properties) yield the same expected revenue for the seller. Thus, the RCN should be indifferent between a second-price sealed bid auction and other commonly known auctions. The logic of a second-price sealed bid auction says that the seller expects the same amount of money as any other auction and the winner gets an object worth more to him than the money. Intuitively speaking, an auction mechanism to allocate sailors would induce an exchange between the RCN and the sailors where the RCN minimizes the cost of personnel and the sailors gain utility from being posted to a preferred location. Thus efficiency gains are achieved. Of course such gains would be vastly superior where the navy is larger and the diversity of naval postings enables finer selection menus for sailors. Golding & Cox (2003) outlines the United States Navy's Assignment Incentive Program (AIP). Several posts within the navy are difficult to fill for many reasons. Previously, involuntary assignment and non-monetary incentives were used to fill these posts. AIP uses a market mechanism (auctions) to allocate sailors and, surprisingly, the implementation of the new mechanism costs roughly a quarter of the previous methods.

Thinking of the future with the Arctic becoming accessible and the RCN facing complicated sovereignty missions with a diversity of postings, this mechanism becomes critically relevant. Thus, should the RCN decide to post sailors in the Arctic onboard the new class of Arctic Offshore Patrol Ships (AOPS), filling such posts will be costly. A program similar to AIP can reduce those costs. Alternatively, it is worth noting that, depending on automation onboard AOPS, many such posts can be replaced with equipment. Most observers discuss aerial robotics applications in the form of drones yet forget that robotics technologies are being applied to naval vessels.

When posts (or the RCN) have preferences over sailors, then an AIP auction may not yield an efficient matching. However, Gale & Shapley (1962) developed what is called the marriage algorithm. To illustrate the algorithm, consider a population with men and women of equal numbers. Each man has a preference ranking over women and each woman has such a ranking over men. Every day, a man proposes to his most preferred woman, and she either accepts or rejects. If a more preferred man comes along and proposes then the woman dumps her current fiancée and pairs up with the new partner. What Gale & Shapley (1962) were able to show is that if unmarried continue to propose until a woman accepts then eventually no new engagements will be announced! Furthermore, the resulting matching is efficient, in the sense that given their ranking within own gender, the man or woman in question cannot improve upon their allocated partner. Matching sailors to posts is slightly different than

the above scenario because there are several sailor candidates for each post, but only one post per man. The problem of the RCN would be referred to as one-to-many matching.

Roth & Sotomayor (1990) provide a thorough review of this literature. The RCN can improve the allocation of sailors by obtaining sailor preferences and the navy's posting preferences over sailors, and by using the appropriate algorithm. Sönmez & Switzer (2012) investigate the effects of similar matching mechanisms on cadet branch or military trade selection. By choosing to serve an extra three years beyond the regular five-year junior officer contract, a cadet may be able to increase her branch choice priority. The Officer Career Satisfaction Program (OCSP) is an incentive program with three alternative returns (graduate school, branch choice and base choice) to the extra three years of service. For example, the selection of graduate study as the benefit is presented with utmost clarity to cadets: "As a cadet, you elected this option to guarantee yourself an opportunity to attend graduate school while in the Army. More importantly, YOU decide whether and when to attend graduate school. Remember: without this option, officers have a one-in-ten chance of attending fully-funded graduate school while in the Army. Also consider that few private sector corporations provide full pay and benefits while their employees pursue higher education. CSP, however, fully funds your attendance at the school of your choice, regardless of cost. If you can gain entry, you can attend, in essence taking a two-year academic sabbatical while remaining on active duty and retaining all associated pay and benefits." (US Army (2013))

The program increased the retention rate of junior officers from 47 percent to 67 percent beyond the eighth year of service. Studies showed that the OCSP option associated with the Gale-Shapley algorithm is mostly responsible for the increase in the retention rate. In terms of human personnel management calculus, the 20% increase in retention of junior officers with eight years of experience but still with moderate compensation relative to officers with twenty years of experience may well offset the cost to the US Army of the investment in officers' graduate education. Moreover, since one of the career incentives in defence forces is the pension obtainable after approximately twenty years of service, the 20% boost to retention also enhances retention by improving the hazard rate, i.e. by adding another three years of service, it enhances the probability that junior officers will serve till the critical twenty years of service.

1.5 Time and Uncertainty

The allocation of resources across time bears significantly on cost effectiveness. For example, suppose the navy planned to have two submarines during peacetime and four during wartime. Building and maintaining submarines takes time and resources. Training regular force sailors is also costly.

One potential solution is to use flexible inputs such as reservists. An obvious cost is in the quality of inputs as reservists are less combat-ready than regular force. The flexibility allowed will outweigh the costs up to a point. Rostek (2006) provides a model of optimal reserve capacity using peak-load theory.

Another solution is to maintain a peacetime regular force that is larger than optimal. This optimal level of reservists and regular force depends on the future stream of demand for security. The future demand for military force is usually unknown as states of war and peace variables cannot be estimated with certainty. Moreover, even in the state of peace, there will be variations of required posturing for deterrence purposes or, in other words, the strategic environment is normally in a state of flux. There are ways to hedge against the risk of war as well as in responding to the evolving strategic environment.

However, whether a navy wishes to hedge against risk or not depends on its beliefs about the evaluation of the strategic environment and how its projected force affects the level of security.

Strategies for reducing exposure to risk are called insurance. In insurance markets, individuals pay a premium and receive a certain amount of compensation should bad outcomes arise. Since insurance markets are seriously incomplete over states of war and peace, non-market instruments exist for insuring oneself. Not only a whole defence force can be seen as a premium against the state of war, but a reservist force can be seen as supplementary insurance against unexpected deviations from the usual strategic environment one of which can be war. There are three ways to summon troops at the outbreak of war: regular forces, reserves, and conscripts. Training conscripts takes time and is often politically infeasible in countries where professional armies have become norm. Regular forces are ideal, but require pay and training during peacetime. Reserves are low-cost during peacetime and an imperfect substitute during war due to lower levels of readiness. Another example of self-insurance consists of assigning soldiers to administrative tasks during peacetime, surely an inefficient matching as they would likely do a worse job than someone who specializes in administrative work. However, these “latent reservists” would be available to assume combat roles when war breaks out as civilian administrators may be hired (or conscripted) quickly.

One’s willingness to purchase insurance depends on several variables. First, the cost of insurance determines its attractiveness. The higher the price the less attractive is the product. Second, defence planners’ beliefs about the likelihoods of different states are important (Grimes & Rolfe (2002)). The more likely the war, the more attractive insurance is. Finally, and less obviously, the manner in which the marginal benefit of security varies with inputs and states is important. If the marginal increase in security increases with military inputs at a decreasing rate then one would call the military risk averse. This is because the expected security without insurance is smaller than the expected security with insurance. It is not enough that the military be risk averse. It must also be the case that the military becomes less risk averse as security increases (Leland (1968)).

2. Substitutions in practice

2.1 Substitutions within navy: Factor and organizational substitutions

Easier in Navy than Army?

The substitution of personnel for equipment with the objective of cost minimization in sight is always more difficult with air force and navy than for army. There are three reasons. First, there is no infantry equivalent in the air force or the navy. These latter two services are essentially platform rather than personnel based although this argument applies to army platforms like tanks. (Pernin et al. (2011)) Armies have traditionally been boots on the ground and continue to be for the foreseeable future. Yet, digital technologies combined with aeronautics have led to the extensive use of drones for many traditional army missions. Second, any platform design comes with fixed input coefficients like a given vessel requires a given number of sailors to run the ship and its combat capabilities. (GAO(2003), Scofield & Brown (2007)) And, finally, the design can hardly be changed through the service lifetime of the platform. This putty-clay technology has been evolving towards flexibility through modularity but

mostly in terms of the output obtainable. A related relaxation of this technological rigidity is at the fleet level where the array of vessels will determine the personnel intensity of the fleet.

The fact that a given class of ships can be only marginally modified over their service life essentially locks in the lifetime personnel costs. One exception may be a more frequent crew rotation during deployment. “The manning of a ship is a major driver of total ownership cost. The Government Accounting Office (GAO) states that “the cost of the ship’s crew is the largest expense incurred over the ship’s lifetime”. This cost is largely determined by decisions made during concept design, which may include significant new support costs ashore.” (Scofield & Brown (2007)) Given the long service life of a navy ship and that a significant substitution of personnel for equipment can only be implemented from one class of ships to the next, a navy’s automation or personnel reduction trend reflects the withdrawal of legacy platforms and their replacement with different and technologically more advanced platforms. Such transitions induce overall decreases in a navy’s overall personnel intensity.

Cost Implications of Crew Rotation

Crew rotation or, at times, sea swap refers to a usual but lumpy substitution possibility in the production of naval defence services. Possible combinations of the two basic inputs, personnel and vessel, enable a given capability to be continuously operational. Knowing that platforms can remain operational for longer periods than crews, two actual possibilities emerge. The first is that a given platform can remain operational by the use of two or more crews rotating to sustain the platform’s longer operational continuity (GAO (2008)). The second, by contrast, is that a given crew is permanently assigned to a platform and, consequently, there are at least two capabilities assuring continuous operation by the rotation of capabilities.

There is actually a third possibility whereby a crew is permanently assigned to a vessel class but not to a particular ship. Consequently, differently than the second case above, economies arise from two sources, one from the vessel’s infrequent journey to home port as it can remain forward stationed between major maintenances, which saves on the ship’s homebound journeys (Hagan et al. (2011)). The other source of economies is that crew redundancy is reduced compared to the first possibility where there may be more crews than ships whereas in this possibility there are as many ships as crews. The crew manning the home port ship can recuperate more flexibly while getting ready for the next forward stationing (Global Security (2012)). This third and the first have been very recently referred to as a major cost saver. “Rather than increase its fleet as desired, the Navy could employ innovative approaches such as “sea swap,” by which some crews are rotated via airplane while ships stay forward deployed longer. Right now, the Navy keeps a single crew on a single ship in most cases, meaning that we waste lots of time sailing ships across oceans to maintain our overseas deployments. This and related ideas could eventually allow the Navy to get by with 260 to 270 ships rather than 286. Ten-year savings could be \$25 billion.” (O’Hanlon (2013))

Assuming all three possibilities are strategically equivalent in the sense that a vessel is permanently forward stationed, the comparative aggregate costs of the options, including non-recurring and recurring costs, ought to determine the selection. In practice, though, the options may have their niches, befitting particular classes of vessels. For instance, the US Navy uses rotational crewing on some submarines, anti-submarine ships and coastal patrol aircraft. It has even successfully experimented with Spruance class destroyers and its successor Arleigh Burke class deployed overseas (CBO (2007)). The new Zumwalt class destroyers have been designed to continue with the crew rotation practice (CBO (2007), O’Rourke (2013)).

Instigated by US Navy's practice of selective rotational crewing, the question arises as to factors determining the optimal scope of crew rotation in the large US Navy. A further question is somewhat presumptive in that whether smaller navies with less forward presence can generalize crew rotation to all their classes of vessels. Notwithstanding the benefits from cost savings and the ability for continuous deployment benefits, crew rotation is not costless.

First and foremost, crew and personnel management issues are more complex with crew rotation than without, especially during transition periods that would characterize new vessels and/or personnel changing vessels. For instance, rotational crewing would require modifications to navy management structures and, in particular, an overall personnel command to oversee all vessels subject to rotational crewing (GAO (2008)).

Second, while the rotation may be straightforward for more general purpose vessels, like frigates, it becomes more problematic with specialized ones as it requires further specific skills and strict team interactions.

Thirdly, and perhaps most importantly, multi-crew ships normally require additional one-time and recurring costs. "One-time costs can be higher because of the need to spend more on designing and building a class of ships for dual-crew use and the need for more-elaborate training and maintenance facilities. Recurring costs can be higher because there are twice as many personnel per ship and because more maintenance is necessary when ships are at sea for longer periods. In return, dual-crewed ships spend far more time on-station in their areas of operations than their single-crewed counterparts do." (O'Rourke (2013)) The one-time costs include design and building costs of further redundancy and ruggedness to ship systems taking into account their longer periods of forward presence. For example, "one-time costs include more-extensive training and maintenance infrastructures. For example, the Navy keeps a larger supply of spare parts on hand (valued at more than \$1 billion) for dual-crewed Trident submarines than for single-crewed attack submarines (whose inventory is valued at less than \$200 million). The dedicated Trident training and maintenance facilities themselves would cost a total of about \$1 billion to replace at both Trident submarine bases. The Trident blue/gold model also incurs higher recurring costs for routine operations. A Trident submarine costs about one-third more to operate than a single-crewed attack submarine. That difference mainly occurs because of the higher personnel costs of having two crews per submarine, but part of the difference reflects higher maintenance costs for Trident submarines. (Excluding personnel costs, operating costs for a dual-crewed ballistic missile submarine are \$46 million per year, compared with \$32 million per year for a single-crewed attack submarine.)" (CBO (2007)) Part of the difference in maintenance costs is attributable to the larger size of the Ohio/Trident nuclear submarines with double (Blue and Gold) crews.

Finally, as a motivational factor, less familiarity with a particular ship reduces the crew's sense of ownership and hence care applied to the platform.

Thus, assuming strategic equivalence between the two crewing options, it's not clear that rotational crewing is a cost-saving innovation in all classes of vessels. It is only a partial analysis to emphasize the cost savings per ship over a given period of time whereas what is required of an options analysis is to integrate all aspects, from a particular ship to all fleet as well as an integration of one-time and recurring costs, including full life-cycle costs of each crewing option (GAO (2008)).

Fleet structure and costs

Cost minimization for a given fleet capability transcends single ship costs, especially in the age of network centric warfare where overall capabilities can be redistributed over component ships of a

fleet(CBO (2006), Johnson & Cebrowski (2005)). For instance, not every ship has to be endowed with frontier technology in situational awareness or in a large spectrum of weaponry as digital networking enables the fleet to behave as a unit delivering those capabilities. “The design of a battle force as a total system, rather than isolated designs of individual ships, can achieve significant improvements in military effectiveness and affordability.” (Bosworth et al. (1991))

There are two dynamic dimensions to changing equipment personnel combinations. The first is the technological innovations that induce automation some aspects of which are covered in section 3.d of this report. The second dimension pertains to the choice of fleet composition. Since some vessels are operated with a higher personnel equipment ratio, strategic choices regarding fleet composition critically affect the size of navy personnel.

For instance, the US Navy’s Ohio-class nuclear submarines (SSBNs) are operated with two crews whereas their sister ships with cruise missiles (SSGNs) have different but still dual-crew rotation arrangements due to the difference in mission lengths. However, the Virginia class attack submarines are operated on a single-crew basis. Amongst the surface combatants, the Spruance class destroyers were operated with 4 crews to a vessel whereas their successor, the Arleigh Burke class, with 3 crews (CBO (2007)).

A reoptimization of the fleet structure through a transition to new classes of ships may facilitate a lowering of the navy life-cycle costs by reducing the personnel-intensity of the fleet. Notwithstanding the debate over reducing this overall cost by lowering personnel intensity versus lowering the number of expensive vessels by crew rotation, capability and cost trade-offs may be inevitable.

Since the required capabilities depend on the strategic environment, possible combinations of classes of ships could be numerous in the case of larger countries (Hoffman (2006)). In fact, the range of vessels in a navy is a strategic asset. “Fleet effectiveness can be enhanced by presenting the enemy with a complex task of having to detect, track, and target large numbers of enemy ships.” (O’Rourke (2007)) Moreover, such a wide range of vessels is an enabler for adapting to changes in the strategic environment. Yet, the downside is economic. The more diverse are the classes of ships, the higher is the overall cost of such a navy as scale economies is traded off for the strategic advantage(Johnson & Cebrowski (2005)).

The US Navy boasts a very wide range of vessels: Ballistic missile submarines (SSBNs), cruise missile submarines (SSGNs), attack submarines (SSNs), aircraft carriers, cruisers and destroyers, littoral combat ships (LCSs), amphibious ships, Maritime Prepositioning Force ships, combat logistics (resupply) ships, command ships, support ships (such as salvage ships and submarine tenders), dedicated mine warfare ships, and sea basing connector ships. Most of these classes being in transition towards the next generation, the reoptimization of the whole fleet is a perpetual process. “As new ships are in development, DOD guidance requires that an analysis of alternatives be completed. These analyses generally include an evaluation of the operational effectiveness and estimated costs of alternatives. In recent surface ship acquisitions, the Navy has not consistently assessed rotational crewing options. In the absence of this, cost-effective force structure assessments are incomplete and the Navy does not have a complete picture of the number of ships it needs to acquire.” (GAO (2008)) Thus cost minimization by fleet reoptimization depends, continuously, on strategic environmental assessments, technology evaluations and, of course, fiscal realities.

Despite the strategic advantage it infuses into a navy by complicating planning and actual operations for future adversaries and by enabling easier adaptation to changing circumstances, a wide range of vessels

also boosts costs the aggregate life-cycle cost of maintaining such a fleet. However, technological advances have brought about possibilities of integrating mission systems smaller ships, of “combining lightweight structural material, innovative design, and sharply reduced manning, a ship’s payload fraction can increase from less than 10% to over 30%.” (Johnson & Cebrowski (2005)) If one adds the improving prospects for modularity into this picture, what emerges is the potential for lowering costs through building smaller ships in longer production runs. Not only smaller ships can act as platforms to existing mission systems but also can serve as diverse ships due to modularity and allow swift insertion of new technologies. Moreover, as already mentioned above, network-centric warfare capabilities further enhance this trend.

Task redesign and automation

A quick historical review of automation and crew reduction shows the secular trend. “Crew size has been shrinking, and the navy has not adapted its maintenance needs to this. This is a trend that has been going on for over a century. In the early 19th century, a typical 3,500 ton ‘ship of the line’ had a crew of 800-900 sailors. That was about 240 sailors per thousand tons of ship. A century later, capital ships had eliminated labor intensive sails and were running on steam, and lots more machines. The 12,000 ton pre-World War I battleship had a crew of 750 (62 sailors per thousand tons of ship). But for the last century, not a lot of progress was made. The current U.S. nuclear carriers have 57 sailors per thousand tons of ship. But the new LCS gets that down to 25. Advances in automation, as well as the introduction of the combat UAVs in the next decade, will make the thousand sailor crew for a carrier possible. That’s ten sailors per thousand tons of ship, plus a lot of robots, and equipment built to require very little manpower to fix or operate. That last innovation is already happening with warplanes, greatly reducing the man hours of maintenance required per flight hour. The navy has long since accepted those concepts for missiles (delivered in sealed containers, requiring little maintenance.) These are trends that have been building for some time, and show every indication of continuing. Although these new techniques are expensive, so are sailors. Each one costs over \$100,000 a year. For a carrier crew of 5,700, that’s over half a billion dollars a year. That buys lots of automation, and keeps a lot of people out of harm’s way.” (Strategy Page (2011)) However, various technological breakthroughs in the past decade or two, from the digital revolution that permitted the network-centric warfare to lightweight materials such as composites, e.g. fiber-reinforced polymers, and robotics all facilitate crew reductions.

A major difficulty with the substitution of personnel for equipment through automation is the mismatch between a narrower range of tasks new equipment can perform than can be performed by personnel it can replace. “... The use of technology and automation is a way to reduce the number of personnel onboard a ship. Technology can be a very effective way to reduce the manning, but it must be used cautiously. Since a single crew member does multiple jobs onboard a ship, there is not a one to one correlation between automating job tasks and removing personnel from the ship’s crew.” (Scofield & Brown (2007)) In a sense, this substitution problem is complicated by the mere fact that the scope economies to bring multiple tasks to be performed by a sailor may not be improved upon by automation. A cost effectiveness analysis would serve as a decision support tool to decide “... whether it is cost-beneficial to retrofit a conventional vessel with advanced technologies that would potentially entail a reduced crew (probably dealing with different and more complex on board duties).” (Lyridis et al. (2005)) Of course, retrofitting is costlier than realizing such substitutions in next generations of vessels even if the retrofit is performed during mid-life refits.

A distinctly pertinent example onboard a naval ship would be the damage controlman (DC). “Damage Controlmen (DC's) do the work necessary for damage control, ship stability, firefighting, fire prevention,

and chemical, biological and radiological (CBR) warfare defense. They also instruct personnel in the methods of damage control and CBR defense, and repair damage control equipment and systems.

The duties performed by DCs include:

- operating, repairing and maintaining installed firefighting systems and equipment, damage control equipment, and chemical, biological and radiological defense equipment;
- training shipboard personnel in the operation, maintenance and repair of damage control systems and equipment, life saving devices, and various firefighting methods;
- performing emergency repairs to decks, structures and hulls by emergency pipe patching, plugging, and shoring;
- performing maintenance and repair of watertight closures and assorted fittings;
- performing emergency repairs to piping fittings and fixtures;
- acting as the ship's Fire Marshal and firefighting leaders;
- training ship's company in chemical, biological and radiological defense." (Powers (2013))

Clearly, the substitution for equipment poses basic problems in complexity. For instance, while the two training systems can be merged into one and partially performed by specialized online training which would replace some DCs, the actual onboard training would still require experienced flesh and bones DCs. Therefore, a critical part of the substitution is the redesign of tasks, perhaps carefully truncating the current set of tasks to make way for equipment that could perform them. "Additionally, while the Navy was reducing the size of its crews, it changed its approach to training sailors and surface warfare officers by replacing some instructor-led training classes with computer-based training. While the Navy has metrics showing that some of these changes have reduced costs and training time, it lacks outcome-based performance measures to determine the effectiveness of the revised training in terms of trainees' job performance, knowledge, skills, or abilities. Without outcome-based performance measures in place to measure the effectiveness of training programs, it is unclear what the effect has been." (GAO (2010)) Similar issues may arise in other tasks. For example, while firefighting can be similarly dealt with, chemical, biological and radiological (CBR) defense equipment might require diverse expert systems that might be too costly to replace specialized personnel with correspondingly wide range of skills.

Another factor in cost minimization is a consideration of risk for human life. Automation, through crew reduction, reduces overall crew exposure to risk (Rohlf's (2011)). Since the cost effectiveness analysis of personnel for equipment substitutions must incorporate costs over the lifecycle of any platform, automation that leads to manpower reduction induces a large reduction of outlays accounting for the risk on human life (Famme (1994)). In general, various "policies", i.e. automation in the present case, must pass a cost-benefit test to be implemented (Viscusi (1993)). However, difficulties associated with an evaluation of benefits may impose simplification towards a cost effectiveness analysis of automation if benefits are comparable in the status quo and automated configurations of vessels in terms of personnel performing a given set of tasks replaced by equipment (Rohlf's (2006), Viscusi (2013)).

2.2 Substitutions within government

This section examines the possibilities of cost savings that could accrue due to cost sharing or scope economies arising from coordinating certain activities within government, in particular through interdepartmental coordination.

Devolution of civilian employment, recruitment and acquisition functions to other federal departments

Shifting certain tasks to other departments would certainly reduce defence department costs but, in the overall budgetary framework, it might not necessarily lead to overall cost savings. However, if scope economies exist in carrying out some tasks or, upon consolidation of tasks in one department, scale economies are generated, the overall governmental cost of performing certain tasks by one department rather than two or more may be reduced, in which case substitutions may reduce costs for all departments concerned.

For example, devolution of civilian employment by defence, recruitment and related services to other departments, of the operation of the supply chain for defence forward bases and depots, and of the acquisition of standard equipment may be one such substitution within government. Another example, to be examined below in Section 2.4 on outsourcing, is the reassignment of SAR from Canadian Forces to Transport Canada or a new civilian agency (Collins (2013), DeMille(2010)).

Supply chain for operational support and forward hubs

In the remainder of this section, two examples will be examined. The first is on forward support bases or hubs, parts of a supply network to support overseas deployments. The 2008 Canada First Defence Strategy anticipates that Canadian Forces will still be deployed overseas in the post-Afghanistan era. In fact, the Canadian government anticipated seven locations around the world on the expectation of future peace support operations. "To date, agreements to establish operational support hubs have been signed with Jamaica and Germany. Each will see a small team of Canadian military personnel stationed permanently within a major, strategically located city served by an international airport and seaport facilities. When needed, the outpost will ramp up and act as a staging area and resupply hub for Canadian missions in the surrounding region." (Berthiaume (2012))

The hub in Germany is expected to cost approximately \$500,000 per year to operate beyond the start-up costs. "However, the benefit to Canada will be huge as the hubs will have the capacity to ramp up operations quickly should the need arise in response to those crises," the then Defence Minister MacKay said, having cited Libya and Haiti as examples of where the Canadian Forces have played a role." (Fitzpatrick (2012))

There are three issues with such forward readiness measures. First, the network of operational support hubs or depots must be optimally built (Girard et al. (2008)). Second, the structure of the hub must take into account the potential missions. Third, the governance of a particular hub must also be optimized, not only in terms of personnel but also of the type of personnel, military or civilian.

When the Canadian Armed Forces establishes a hub in a potential operational theatre, the hub at cadre rather than caretaker status has limited facilities and a small staff up to 10 personnel, sufficient to deliver essential operational support possibly including warehousing of pre-positioned materiel while preparing for potential full activation (Canadian Forces (2013)). The essential problem related to the core of this study is that, when the hub is at caretaker status, it does not need military personnel. The

question arising is, then, whether the Canadian Foreign Affairs or International Trade missions can operate it at a lower incremental cost than the operational cost of a standalone Defence facility.

The second example is the joint U.S. and U.K. forward support base at Diego Garcia. This joint base is situated on a tiny atoll about 1,000 miles from India and tasked with providing logistical support to the two countries' forces when they are deployed in the region. During the wars in Afghanistan and Iraq, the base was crucial. However, given that it is ultra-remote, shipping materials and maintaining the facility itself can be highly costly and, in addition, the substitution suggested for the Canadian forward depots is simply infeasible. Yet, beyond its strategic military location, Diego Garcia benefits from its remoteness to be a key observatory for tracking satellites, it is one of five monitoring stations for GPS and, additionally, the island is one of only a handful of locations equipped with a Ground-based Electro-Optical Deep Space Surveillance system for tracking objects in deep space (Sweeney (2010)). Its relevance beyond defence reduces the cost of the military base by the multiplicity of scientific outputs produced by the island facilities.

2.3 Substitutions by ship design

Cost minimization by input substitution is often perceived as a process that succeeds the design and development of a product or service. However, although design is inextricably related to the inputs used in the production process, the outcomes expected from the product dictate what is being designed and a large percentage of the total cost is determined at the design stage (Caprace (2010)). Moreover, design fast becomes complex with more capabilities expected from the product. For instance, though expensive, personnel are typically capable of multi-tasking whereas, as substitute, a given piece of equipment may not perform multiple tasks. Nowhere in shipbuilding has this requirement come to the fore than in damage control that requires intensive multi-tasking. This in fact proves to be a scope multiplier in the sense that functions that could have been performed by different equipment can better be performed by multi-tasking personnel. Given the available technologies, there exists a lower bound to vessel crews. Nowhere this lower bound is more evident than in areas of damage control and multi-mission simultaneity. In particular, ship design has to take into account the "reliability of autonomic responses to damage events and the crew situation awareness and decision making in the face of the immense stress and pressure that accompany fire, flooding, or other casualties." (Hagan et al. (2011))

This section focuses specifically on design features that would allow the substitution of personnel for other inputs and equipment.

Design and Development Cost

A switch from a large and complex thus expensive to a lower-cost design could induce substantial design and development cost savings. It might also save money over the long run if the lower-cost design is produced in so large numbers that the cumulative procurement savings were greater than the additional up-front design and development costs. Thus, considering a complex project under way, if the switch is contemplated early and the restart button is pressed quickly, the lower-cost alternative design might reach the break-even point earlier for overall cost savings. In addition, if a lower-cost design could use many of the same technologies intended for the more-expensive ship, or technologies already developed for other ships, then the cost to develop the new design could be reduced, perhaps substantially. Since newer designs incorporate more automated functions, the switch is normally a labour-saving innovation through a smaller crew. Of course, as considered in the theory section of the

report, procurement costs are incurred in the short run whereas savings realized from lower personnel costs are achieved over the service-life of the platform.

A design durability issue arises particularly with surface combatants when a vessel is retired before its designed service life. Such durability is increased with modular designs that allow mission systems flexibility. However, a design process change is as relevant to this analysis as the design change towards modularity which, by itself, might enable substitution toward more capital-intensive vessel designs. A new design technology whereby platform and mission systems are decoupled and designed concurrently rather than sequentially allows a mission system five to seven years younger than the current technology used (Koenig (2009)). Assuming that, with the applications of fast-advancing frontier technologies, some onboard ship capabilities can be delivered by more equipment-intensive technologies, this modification of the design process can deliver more labour savings.

Modular Adaptable Ships

There are two versions of modular adaptable ships by design. The first is an adaptable ship by design in the sense that the multi-mission vessel can adapt to the strategic environment by modifying its available mission modules. Cost savings will accrue because the client navy can build its target capabilities through a smaller number of ships. The second uses the commonality of certain parts of the vessel during construction but the end product may be a single-mission vessel once sailing. Thus, in this second case, savings accrue in the construction phase. The US Littoral Combat Ships (LCSs) and the British Type 26 Global Combat Ships (GCSs) (Stoker (2012)) are the precursors to modular adaptable warships that are currently entering service.

When the Freedom and Independence classes of American LCSs were planned in late 1990s, designers envisioned quick reconfiguration of combat capabilities by swapping mission modules in the heat of naval combat. For example, the anti-submarine and surface warfare mission modules could be interchanged in the presence of an evolving threat and such module changes required significant automation if to be performed within hours. Beyond the littoral protection missions, these “plug-and-fight” mission modules provide, in theory, significant combat flexibility within a single hull and cost savings in terms of having to maintain a smaller number of ships. “In contrast to the traditional approach, which is to cram a wide-ranging set of bolted-in compromise equipment into fixed installations, “flex ships” can radically change the ships’ capabilities, by swapping in a full breadth of equipment focused on a particular need.” (Defense Industry Daily (2013a)) However, given the current technologies available, switches took weeks in simulations. Such lessening of combat advantages has now delayed the introduction of multi-mission ships in the near future (Atherton (2013)).

An example to the second class of modular adaptable ships is the modular hull ships. A common bow and stern may be used for several classes of ships whereas a parallel midbody may house various mission systems and equipment that may not only be replaced frequently over the service life of the ship but even within the same mission. This may reduce the scheduled and unscheduled maintenance down and retrofit times. The first class of cost savings includes, as a platform for multi-mission capabilities, the specifically designed mission bays. Continuing with LCSs, both LCS variants include a mission bay to house mission packages, which are typically composed of mission modules, aircraft, and crew detachments to support the mission modules and aircraft. Mission modules in turn are composed of mission systems and support equipment. The former includes weapons, sensors, and vehicles whereas the support equipment consists of support containers, communications systems, and computing environment (Ailes (2013), Doerry (2012)).

The wide range of mission packages available for LCSs includes mine warfare, surface warfare, anti-submarine warfare, humanitarian assistance/disaster relief, counter-piracy/maritime security operations, harbour defence, immigration control/border protection, emergency medical, Vertical Takeoff Unmanned Air Vehicle (VTUAV) support, naval aviation support, electronic warfare, irregular warfare, and special operations (Covert (2012)). Without underestimating the engineering challenges in the hull and package designs, and in system integration, this rather wide variety of packages certainly enables shipbuilders to implement cost-saving measures of the second category. Moreover, cost savings of the first category can be implemented by navies once the swapping technology is developed to make module switches sufficiently swift.

Although recent classes of US navy ships are not available to other countries, the German shipbuilder Blohm + Voss of Germany, a subsidiary of ThyssenKrupp, incorporated modularity into their MEKO surface combatants and enabled Blohm + Voss to rapidly and affordably create customized warship designs for domestic and foreign military sales using standard components. Blohm + Voss sold over sixty MEKO vessels in over 15 configurations to many countries. Moreover, Howaldtswerke-Deutsche Werft (HDW), another subsidiary ThyssenKrupp that builds submarines, integrated modularity into their U212 and U214 class submarines (Naval Technology (2013)). Similarly, benefiting from modularity in design, Damen Schelde of the Netherlands developed small vessels with their SIGMA technology and exported internationally, both naval warships and civilian vessels (Gelling & Goossens (2008), Damen Schelde (2013)). Yet another shipbuilder in the fold is Spain's Navantia with its Buque de Accion Maritima (BAM) class patrol vessel built on automated platform which allows different configurations supporting various mission requirements (Naval Technology (2012)).

The third category of cost savings arises from scale economies achieved by modularity technologies. As demonstrated by Damen Schelde's small vessels example, the increase in production runs due to hulls usable as civilian vessels generates scale economies (Gelling & Goossens (2008)).

Yet a fourth category of cost savings is due to evolutionary acquisition properties enabled by modularity. The open system architecture supports the future insertion of new technologies as packages into the system of systems that is the modularity. Since no fundamental redesign is necessitated whenever the new packages are designed to function with the existing system, major cost savings can be achieved (Volkert, Jackson & Whitfield (2010)).

One result of the evolutionary acquisition of new mission packages and their system integration under modularity is that the vessel's service life becomes endogenous not so much for the packages but rather due to the tradeoff between an old hull and a new one that has been designed to facilitate the integration of current and future packages. "As modular weapon systems allow cost-effective upgrades of a vessel's war-fighting capability, the degradation of the difficult-to-upgrade structure of the vessel may soon become one of the key drivers of vessel retirement and lifecycle maintenance costing." (Collette (2011)) With the current developments in railgun and directed energy weapons technologies, a decoupling of ship construction and weapon system development is a reality. No design and platform changes may be necessary for retrofitting modules into in-service vessels. A downside of this modular adaptability is inaccuracies arising in the estimation of acquisition or, in particular, the life cycle cost of naval systems (Doerry (2012)). Updating of radar systems is an example to this endogeneity. "The topside arrangements of all the Radio Frequency (RF) transmit and receive antennas is a challenging task. Ensuring electromagnetic compatibility (EMC) while minimizing electromagnetic interference (EMI) and antenna blockages is difficult even with a fixed set of known RF equipment. Over the service life of a ship however, these RF equipment may require replacement or upgrading to remain interoperable with the fleet and militarily relevant. Currently, replacement and modification of RF equipment and their

associated antennas are not extensively considered or accounted for in shipboard topside design. Upgrading arrays and antennas can be extremely expensive. In particular, phased array radars have traditionally been tightly integrated into the ship superstructure design. When these radars become obsolete, the cost of modernization may drive a decision to decommission the ship prior to its designed service life rather than invest in updating the radar.” (O’Rourke (2005)) With open architecture and modular adaptability, major refits and maintenance, both labour intensive activities, appear to become redundant in new generations of vessels.

Despite current technological difficulties in perfecting the modular adaptable ships, there exist strong reasons why the evolution will proceed. First, smaller fleets based on multi-mission capable vessels may reduce overall costs. Although design and procurement costs may increase, lifetime costs of operation and maintenance will fall by saving on personnel and the number of ships required. Second, given faster rates of technology change, especially in software, modularity will reduce the number of costly refits and maintenance (Abbott (2006)).

Procurement Cost

Related to design but realized at the construction phase, vessels can be designed in larger hence smaller numbers of modules to be assembled. This reduction in construction complexity can shorten the assembly period, lower the use of skilled labour and hence lead to lower overall procurement costs. An example is the Virginia class attack submarines where the assembly period was reduced from 72 to 60 months (Johnson (2009)).

It is well known that moving down the learning curve is faster the higher the number of ships built. As a shipyard builds more ships to one particular design, the shipyard production team becomes increasingly familiar with the design and implements cost-cutting innovations provided incentives exist. If a simple design of lower-cost can be built at a higher annual rate than a design under way, then the lower-cost design could move down the learning curve more quickly and exploit the cost-reducing benefits of the learning curve more fully than the incumbent design. Furthermore, if the simpler design incorporates onboard-ship labour-saving substitution, such a design change not only induces fleet adjustments but also, by virtue of less labour-intensive ships, further substitution of labour for automation. The next section provides an example.

A fourth option would be to change the acquisition strategy for building certain Navy ships. For example, changing from the current strategy of building each Virginia (SSN-774) class attack submarine jointly by two yards to a strategy of using a single yard to build all Virginia-class boats might eventually reduce the cost of each boat by roughly \$60 million to \$180 million. As another example, the Navy estimates that changing from the currently planned strategy of dividing DD(X) destroyers evenly between two yards to a strategy of having all DD(X)s built by a single yard could reduce the cost for building 10 DD(X)s by a total of \$3 billion, or an average of \$300 million per ship. In either case, however, shifting to a single yard acquisition strategy could cause the second yard to permanently exit the business of building that kind of ship. (Thus an option value cost!) That could leave the Navy with a single source for building that kind of ship, which could prevent the Navy in the future from using competition or benchmarking to spur design innovation, constrain costs, maintain production quality, and ensure adherence to scheduled delivery dates.

A fifth option would be to take steps to reduce the amount of shipyard fixed overhead costs that are incorporated into the procurement costs of Navy ships. This could be accomplished by eliminating any excess capacity among the yards building Navy ships, which would eliminate the fixed overhead costs

associated with maintaining that capacity, or by increasing other kinds of work done by those yards, so that this other work could absorb a greater portion of the yards' fixed overhead costs. Potential other forms of work include construction of ships for other U.S. government agencies, such as the Coast Guard or the National Oceanic and Atmospheric Administration (NOAA), construction of commercial ships, overhaul and repair of Navy or other U.S. government ships, and overhaul and repair of commercial ships. By reducing the significant overhead burden arising from maintaining expensive platforms, the British air-to-air refueling aircraft/strategic transport option, where the contractor can do private work, offers a solution with the drawback being the potential hazards of contractual sharing of peak-load or surge priority right to services (NAO (2010)).

Personnel and Maintenance Costs

A ship's life-cycle O&S cost can be reduced at the cost of increasing its procurement cost. Crew size can be reduced by ex-ante automation of functions hitherto performed by personnel. For example, damage control is a function that requires a larger crew size than otherwise. If the cost of automation technology, inclusive of skilled crew costs, is lower than the cost of damage control personnel then substitution of personnel for automation will be desirable. Interestingly, the cost of damage control personnel includes, beyond compensation for personnel, designing and building crew-related spaces into the ship. Moreover, maintenance generates another major component of a ship's O&S cost. Reducing maintenance costs might require building the ship with more-durable but more-expensive materials, or increasing the size (and thus construction cost) of certain spaces on the ship, so as to provide room for easier access during maintenance (O'Rourke (2005)). Therefore, the ship design has to trade a lower procurement cost off against a higher life-cycle O&S cost. Finally, as a last example, advanced bottom paints⁴ will reduce the frequency of re-painting, a labour-intensive maintenance activity.

Fuel Costs

Different navy vessels exhibit different power requirements, typically the larger the vessel, the higher the power requirements. There is, however, a significant difference in terms of manning requirements between nuclear and conventionally powered vessels. For most vessels except aircraft carriers, a nuclear power plant may last for the entire life span of the platform whereas, for conventionally powered vessels, there will be steady O&S costs not only due to fuel replenishment but also to maintenance of the power plant. Since the nature of the capability matters, there may well exist other relevant criteria for power plant selection. For instance, since nuclear powered submarines are only constrained by onboard food stocks, they are particularly suitable for missions through the Arctic. Otherwise, if selection depends on life cycle costs, nuclear powered vessels typically have higher procurement costs and lower O&S costs, their aggregate costs depending on particular ships. Of course, apart from this latter case, selection of power plant critically depends on expected oil costs (O'Rourke (2005)).

2.4 Substitutions by outsourcing aspects of military operations

Outsourcing is organizational substitution in the sense that tasks performed by part of a larger organization are now performed by another. For example, in the field of the civilian aircraft construction, no less than eight countries' manufacturers supplied parts for Boeing's Dreamliner which has been the subject of intense debates over the scope of outsourcing, all in the private sector

⁴ Of course, not all paints are neutral to operational effectiveness. Magnetic signature must be factored into operational effectiveness consideration.

(Fernholtz (2013)). This report concentrates specifically on outsourcing tasks from a defence organization to private sector companies. Similar issues persist but the main motivation is cost savings. Logistics and procurement have been primary areas where national defence organizations have traditionally incorporated outsourcing but, lately, outsourcing has been extended to combat operations (Mobley (2004)). Outsourcing can work in a variety of ways to lower costs: Lower personnel use through automation by contractor; lower levels of contingency personnel in the mother organization; transformation of fixed costs into variable costs at the premium of insuring PMCs for the variations in demand; contracting with locals in deployed operations rather than using home country resources.

Legal Considerations

There are potentially many legal limitations on privatizing parts of national defence. Barriers in national legislation should be considered as minor hurdles because, in the Westminster system of parliamentary democracy, parliamentary supremacy allows any elected government with majority to impose a legal framework through legislation that permits the use of private resources in military services provision. However, there are clauses in some agreements in international humanitarian law (IHL) which restrict the use of private military resources by sovereign nations, even if they are not actually used in combat. We examine these limitations and their applicability to Canada below.

We begin by examining the legal status of PMCs and their employees themselves. It is clear that, from a practical standpoint, the line between soldier and non-combatant is at times difficult to draw during armed conflict. In a Green Paper written for the United Kingdom's House of Commons (UK Green Paper (2002)) about the potential regulation of PMCs, it is asserted that "The people who fly soldiers and equipment to the battlefield are as much part of the military operation as those who do the shooting. At one remove, the same applies to those who help with maintenance, training, intelligence, planning and organisation – each of these can make a vital contribution to war fighting capability" (O'Brien (2002)). Consider the common occurrence where a PMC is contracted by a government to engage in a purely defensive operation such as providing site or convoy protection. In fact, the US military frequently employed Blackwater USA (now known as Academi), a PMC, to protect convoys and personnel during the Iraq War. If, in such circumstances, the asset came under attack, necessitating PMC personnel to engage in direct conflict with the enemy, how would the PMC members be classified? Indeed, some legal scholars see their status as a legal gray area (Singer, 2005). Attempting to answer these sorts of questions from a legal standpoint, we begin by examining several definitions outlined in the Geneva Conventions (GC), to which Canada is a signatory. The two relevant definitions are *combatant* and *mercenary*. GC defines combatants under Article 43: "The armed forces of a Party to a conflict consist of all organized armed forces, groups and units which are under a command responsible to that Party for the conduct of its subordinates Members of the armed forces of a Party to a conflict (other than medical personnel and chaplains covered by Article 33 of the Third Convention) are *combatants*, that is to say, they have the right to participate directly in hostilities." Thus, even military personnel with roles far removed from front-line combat, such as cooks, have the right to engage the enemy in combat and, as a result, also have other rights conferred upon them such as the right to be a prisoner of war (GC Article 44). If PMCs under the employ of a nation are not directly under its military chain of command, it can be interpreted that they do not have an explicit right outlined in GC to participate directly in the hostilities.

The point of contention here is the use of the word "directly". There is an ongoing debate as to what is meant by the word: "The current controversy discusses whether the protection of civilians against unlawful attacks from a party to the conflict or the provision of security of military buildings or targets

against irregular combatants can be considered a direct participation in hostilities.” (Rousseau, 2012) Other confusions abound. By the letter of the law, persons performing support roles (excluding espionage – see GC Article 46) at the behest of their patron state but outside of the nation's military chain of command are considered non-combatants (that is, civilians). This would seem beneficial, as civilians are afforded considerable protection under GC. However, contractors performing such functions may be seen by the enemy as de facto members of their nation's military and thus legitimate targets of warfare.

Given the difficulty in defining where various PMC activities fall on the spectrum of combat involvement and for reasons of space, we simplify the analysis to assume that Canada's military would like to employ PMCs to participate in direct combat; clearly, if such an activity is permissible under international law, so would any activity to support it.

To answer this question, we turn to the issue of mercenaries. There are 6 conditions that all must be satisfied for a person to be classified as a mercenary. According to GC Article 47 subsection 2, a mercenary is any person who: “(a) is specially recruited locally or abroad in order to fight in an armed conflict; (b) does, in fact, take a direct part in the hostilities; (c) is motivated to take part in the hostilities essentially by the desire for private gain and, in fact, is promised, by or on behalf of a Party to the conflict, material compensation substantially in excess of that promised or paid to combatants of similar ranks and functions in the armed forces of that Party; (d) is neither a national of a Party to the conflict nor a resident of territory controlled by a Party to the conflict; (e) is not a member of the armed forces of a Party to the conflict; and (f) has not been sent by a State which is not a Party to the conflict on official duty as a member of its armed forces.”

There are four issues concerning application of Article 47. The first is that the definition of mercenary in GC is quite restrictive; in particular, clause (c) is especially problematic since it is hard to prove: it is far from trivial to demarcate what compensation level would be “substantially in excess” of normal, and it is likely difficult to establish motivation for private gain in a cogent fashion (Milliard (2003)). The second concerns its applicability: it can be argued that the employment of mercenaries is not prohibited by the Article; rather, it only serves to discourage their use, as they are not considered lawful combatants and are thus not entitled to prisoner of war status if captured by the enemy (Milliard (2003)). Mercenaries are still granted protection if captured: Article 45 subsection 3 outlines their protection. Use of mercenaries is not considered a grave breach of GC, and is thus not considered a war crime. Grave breaches are outlined in Article 85. The third issue is that it is generally thought that Article 47 does not cover soldiers employed by PMCs, i.e. PMC employees are not mercenaries, regardless of their role. In 2005, a meeting of experts discussed, inter alia, as to whether a PMC employee could be considered a mercenary under GC. The consensus was that PMC members do not fall under this definition (Gifford (2008)). Some scholarly work echoes this view (e.g. Cameron (2006)). This interpretation is buttressed by the perception that Article 47 was motivated primarily due to the mercenary trade in the war-torn countries of Africa (Milliard (2003)). The fourth concern is that it is possible for a nation to circumvent Article 43 through legislative or executive acts designed to bypass one of its requirements, e.g. by granting PMC employees official status as members of the armed forces. Such an end run around GC has been attempted in the past: for example, Sandline International, a British PMC, in a contract with Papua New Guinea in 1997 which ultimately did not come to fruition, employees were to be termed “Special Constables”, sidestepping the mercenary definition by no longer having PMC members satisfy requirement (e) of GC Article 47.

There are other international agreements forbidding the use of mercenaries, such as the International Convention against the Recruitment, Use, Financing and Training of Mercenaries (referred to informally

as the UN Mercenary Convention). This agreement provides a slightly broader definition of mercenary than GC, but Canada is not a signatory and is thus not legally bound by it. It must be underlined that most signatories are African countries. The United States as well as most Western European countries (including the United Kingdom) are not a party to the treaty either.

Two nearly century-old international agreements have clauses which can be interpreted as restricting the use of mercenaries, but Canada is not a signatory of these either. The two agreements are (i) the 1928 Protocol to the Convention on Duties and Rights of States in the Event of Civil Strife, and (ii) the Convention concerning the Rights and Duties of Neutral Powers in Naval War (signed in The Hague, October 18th, 1907). Another avenue of consequence is the UN Security Council, which through a declaration could impose sanctions or other harsh penalties on Canada for its use of PMCs in combat. However, given the veto power of three of the permanent members of the voting members which are close allies of Canada (specifically, the United States, the United Kingdom, and France), and the fact that the first of the three uses PMC resources extensively in controversial situations, a scenario where Canada would face penalties for such actions appears extremely unlikely.

The last issue we consider is the prosecution of individual officials under international law for the use of mercenaries. The International Criminal Court (ICC), beginning in 2017, will have the power to prosecute individuals for the crime of aggression, defined as “the use of armed force by a State”. However, the amendments are restrictive in that their definition of aggression excludes the use of force by non-state entities, a category in which mercenaries fall under (Gillet (2013)). Therefore, the threat of prosecution under this statute appears minimal.

To conclude, other than through an ex post facto basis, countries do not appear to be subject to non-trivial legal repercussions for the use of PMCs for frontline combat operations, ipso facto. Therefore, all other uses of PMCs to support combat are likely permissible without legal consequence in international law. There are numerous other legal considerations concerning the employment of PMCs, depending on their role. Their use in combat operations, even defensive, carries the most significant risks. Nations are in general held accountable (in various fashions) for the behaviour of their troops, subcontracted or otherwise. Since PMC soldiers are typically not under the direct military chain of command, their behaviour on the battlefield may be more difficult to police, whether or not there are accompanying regular forces. There is a clear incentive problem: soldiers under the employ of a PMC may have a lot less to fear legally compared to a soldier that fights for a nation's army. This is especially true in the case where the contracted PMC is not based in the contracting nation or in any of its close allies. Militaries may attempt to curtail undesirable activity by contracting against it – for example, PMC employees who accompany a US armed force in the field are subject to the Uniform Code of Military Justice (UCMJ) in its entirety. Despite holding contractors accountable to military law (or other standards of behaviour), imperfect monitoring may increase the likelihood for contractors to behave: if using a private military force with little supervision on the ground, it may be very difficult to hold them accountable for their actions. These problems, while endemic to warfare, extend well beyond the battlefield: the lack of supervision of the subcontracted interrogators at Abu Ghraib prison in Iraq was thought to be a contributing factor to the abuses perpetrated there (GAO (2005)). As we will see, potential problems due to imperfect monitoring stretch well beyond the misbehaviour of PMC employees.

Market Imperfections

There are many imperfections in the market for defense services provision. The chief problem is asymmetric information; while there may be issues due to the relatively small number of defense contractors to be discussed below, these market power problems are of a lower order of concern

because the military can provide its own services if outside options are unfavourable. We discuss these various imperfections and their implications for defence contracting below.

The primary issue that the government is faced with when contracting out defence services is the principal-agent problem. Most information asymmetries arise after the procurement contract has been signed, and the objectives of PMCs and those of the governments that employ them are not necessarily aligned. Therefore, there is serious potential for moral hazard (Mas-Colell et al. (1995)). While governments are interested in achieving policy goals, the primary objective of a private company is to maximize its net present value, which roughly translates to maximizing profits. Of course, such situations of possibly divergent goals being a normal feature of two actors working together, this sort of problem may be more serious in the context of military outsourcing.

Consider the current and desired status of a conflict from the point of view of an intervening government (which has the policy goal of ending the said conflict) and that of a PMC working for the government. Since the contracted PMCs benefit as long as the conflict is ongoing, they have an incentive to act in a manner that extends its duration. Of course, this is probity concerns in the transaction costs literature (Williamson (1999)). In the literature on “war economies”, the absence of order creates economic benefits for some parties, who may have little interest in moving to a state of peace (Duffield (1994)). An example of such an economy would be Somalia, which has been in an almost perpetual state of civil war since 1991. There, various factions have benefited financially from the continuous state of violence, and it has been predicted that there will be continued local resistance to international efforts to end the violence there (Shortland et al. (2013)). While the possibility that a contracted company may not act in a manner consistent with ending a conflict as soon as possible may seem far-fetched, less-reputable parties have been contracted in the past to provide security; for example, in Somalia, humanitarian organizations have contracted local warlords to provide them with protection (Singer (2005)). Of course, contracting with reputable organizations will do much to assuage this concern.

The usual primary motivation to contract the services of PMCs is to benefit from cost savings. However, depending on the role, it may be difficult to ascertain whether such a savings is taking place, because it may be difficult to compare the performance of the contracted company with the status quo (or with other PMCs). To illustrate this point, consider the case where a car company would like to outsource production of an engine component. It can outsource it to several companies. If dissatisfied with one, it can hire another, and it can easily compare the performance of both contractors with each other and with itself (e.g. by seeing the average cost per part, by gauging the quality of the parts, etc.). This is normally not possible with some kinds of PMC contracts, since many situations are one-shot, e.g. one cannot compare how firm A would have done in the situation when one hired firm B. As an example of this, consider instead the use of a PMC to perform some combat objective, such as to protect a supply base and its nearby civilian population from attack by the enemy. How does one compare the performance of one firm to another, or compared to how the national military would have performed? One would need to make a number of assumptions, some potentially very strong, to gauge the performance of the party that completes the contract in order to compare it to how others would have performed. Strong assumptions may even need to be required when a similar situation re-presents itself in a future period: for example, people often cite the swift victory of Executive Outcomes (“EO”, a PMC based in Africa) over the Revolutionary United Front (“RUF”, a rebel force) in Sierra Leone in 1995 as proof of their efficiency, as it took the UN several years to defeat the RUF when the war restarted after EO's contract was terminated (Singer (2005)). To come to such a conclusion, one would have to assume that the conditions EO faced in 1995 were exactly the same as the UN faced when they were sent to Sierra Leone in 1999: that the RUF had the same or similar number of troops, access to infrastructure, used similar tactics, and possibly other conditions. Most would view these as strong assumptions.

In some environments, proper surveillance to ensure that contractors are working efficiently is likely required. Clearly, contracts based on hourly wages at fixed hourly rates require outside supervision, because there is no incentive for the producer to complete the work as quickly as possible. However, this concern should be weighed against the fact that soldiers also have strong incentives to shirk. A soldier who completes an assigned task ahead of the time allocated to do it may be sent to work on something else; if said soldier does not believe that the benefit of finishing ahead of time (e.g. to demonstrate competence and thus increase the probability of promotion) exceeds the cost of effort, they will not have an incentive to finish early.

Adequate monitoring is not only required for employees, but for the procurement of goods promised by the contracting company as well. As an example of how much money could be lost, during the Iraq War, Halliburton charged the US Department of Defense (DoD) \$247 million dollars in goods that were not delivered. This total consisted of \$186 million in meals not served, and \$61 million on gasoline that was not provided. There was an extensive legal battle between the US Army and Halliburton over the unaccounted charges. (Minow (2005)). The contract's value where this wrongdoing occurred was \$10 billion; hence, roughly 2.5% of the contract's value was potentially lost.

When faced with a principal-agent problem, the principal (in this case, the government) may seek to write a contract in such a way to induce the agent to behave in a fashion that is consistent with completing the former's objective. However, the uncertainties associated with warfare, along with the structure of the market for military procurement, can make writing a contract that effectively serves the interests of the government quite difficult, depending on the role of the contractor.

The optimal design and enforcement of contracts has been a serious issue in previous experiences of military procurement. The United States, which made extensive use of subcontracting during the Iraq War⁹, has experienced significant problems enforcing contracts due to failures in oversight (Minow (2005)). For the fiscal years 2005 to 2010, the US Department of Defense (DoD) appropriated \$112.1 billion dollars for contracts in Iraq. This amount totals approximately 19% of total DoD spending for the area (Schwartz & Swain (2011)). Some contractors purposely evade monitoring intended to evaluate their contractual performance (Guttman (2004)). Not all problems related to contracting are on the side of the contractor: for example, at the Abu Ghraib prison in Iraq, military personnel did not know basic information about contracted personnel, such as the terms of their contract, the procedures, or even guidance as to how to use them (Schooner (2005)). We outline the various contracting-related issues below.

One significant obstacle in designing an efficient contract, i.e. a contract that can only increase one party's benefit at the expense of the other, is the fact that, due to the uncertainties associated with war and conflict, contracts are sometimes quite incomplete, i.e. some scenarios or events are not accounted for. A complete contract is something that only exists in theory, as there are always events that can occur which are deliberately left out (Halonen-Akatwijuka & Hart (2013)) or are insufficiently covered by the contract. Depending on the task, some contracts may be more incomplete than others. For example, subcontracting a PMC to take over most of the recruitment process of a nation's military entails far less uncertainty about costs due to unforeseen developments compared to subcontracting a PMC to create a humanitarian corridor to protect a local population from genocide. One key issue is the assignment of residual rights, i.e. whenever an unforeseen event occurs, a clear assignment of residual rights allocates decision-making to a particular party to the contract.

Continuing with the latter example, suppose that, when the PMC attempts to establish the corridor, it finds that it has insufficient military power to hold it against the belligerents. Does the PMC have to ask

the contracting government for more money first, or can it just immediately deploy additional forces and charge the contracting government for it? These sorts of issues need to be taken into account when writing the contract. In fact, the more incomplete the contract, the more the PMCs will benefit, regardless of who owns the residual rights. In our example above where we discussed a contracted PMC establishing a humanitarian corridor but needed additional military forces, consider the following two cases. In the first, the PMC can deploy additional forces to secure the area without the contracting government's approval. If it has this right, there is no incentive for the PMC to do so in a cost-efficient manner, so it may deploy more forces than are needed in order to obtain additional profit. In the second case, the government must first give approval to the PMC before it can deploy additional forces. Here, the PMC can take advantage of a sort of "hold-up problem", where it has additional bargaining power compared to the position it was before it signed the contract. For example, the PMC can threaten to withdraw from the area it was contracted to protect if not granted additional funds because, by defending an area larger than it can manage, its soldiers would be faced with a level of danger that far exceeds what was contractually agreed to (either implicitly or explicitly); hence, the PMC could bargain for more profit.

Clearly, it may be (Halonen-Akatwijuka & Hart (2013)) in the government's best interest to design a contract that is as complete as possible, and to use historical experience to contract the appropriate response to reasonably likely contingencies. When insourcing such tasks, these two contractual hazard problems are not present.

Implicit in this discussion is the question of switching costs. The bargaining power of a PMC may increase after the contract is signed due to these costs: in the example above, the cost to switch to another PMC (or to regular forces) during a humanitarian crisis is potentially very high, both in terms of dollars and in terms of human lives. For mundane activities, switching costs are potentially much lower, and thus better deals can be negotiated with PMCs. It should be noted that there is much potential for inefficiencies when subcontracting PMCs, especially if the contract is unspecific. One consideration is that PMCs can further outsource activities that they contractually agreed to perform: for example, in Iraq, Halliburton subcontracted out work that it was contracted to perform, so at times there were 2 or 3 layers of subcontracts used in the completion of a given task (Minow (2005)). These multiple layers of contracts can be thought of as analogous to "red tape" problems associated with government bureaucracy.

The one factor that counterbalances the problems with incomplete contracts is reputation. Returning to the example of the humanitarian corridor, consider both cases once again. In the first case where the PMC can deploy additional resources without permission, it may do so with restraint, knowing that if it does so in an inefficient manner and is detected doing so, it decreases the likelihood that the government will do business with the firm again. In the second case where the hold-up problem was present, a firm concerned about its reputation would not make excess use of its increased bargaining power as doing so would also reduce the likelihood of obtaining additional contracts. Under both scenarios, reputation serves to reduce the potential problem.

While the problems discussed above are present no matter what the size of the market, the issue of the competitiveness of the market is especially relevant here. The number of competitors in the markets for various PMC services can be quite small, especially in frontline combat roles, where there are only a few players. In the US, there also appears to be a concern over the number of no-bid contracts (Guttman (2004)); for example, the government argued that only Halliburton could feasibly complete the contracted duties it asked for in Iraq (Minow (2005)). The US government has concluded that steps need to be taken to increase competition in defense contracting for both goods and services (GAO (2013)).

The market for defense services is clearly a situation of few buyers and few sellers, so the market may not be competitive. A market may still be competitive despite the small number of players; for example, the large passenger airplane market is normally seen as competitive despite the presence of only 2 major players, Airbus and Boeing. If there is market power in defense services provision, potential savings from subcontracting may be lower as PMCs may implicitly collude on pricing.

A Simple Model of Outsourcing

The decision of whether to outsource can be succinctly illustrated in the following game. First, the military(M) chooses whether or not to subcontract. Then, the state of nature is determined: either things go according to plan and no unforeseen incidents occur, or the situation goes awry. If the PMC is employed, it can take advantage of the contract incompleteness. For the ordered pair (a,b), let “a” be the payoff to the military and b be the payoff the PMC. As these are payoffs in terms of utility values, higher is better. The extensive form game is thus as follows.

Military moves first:

- Military chooses to subcontract (action S)
 - Nature moves
 - Things go as expected (action E): Payoffs (12,5)
 - An unforeseen incident occurs (action U): Payoffs (0,10)
- Military chooses not to subcontract (action I)
 - Nature moves
 - Things go as expected (action E): Payoffs (10,0)
 - An unforeseen incident occurs (action U): Payoffs (5,0)

The payoffs in this game are so chosen as to illustrate the following scenario: if the military subcontracts and things go as expected, they gain (as the payoff of 12 for {S,E} exceeds the payoff of 10 for {I,E}). However, if the military subcontracts and an unforeseen incident occurs, they lose (as the payoff of 0 for {S,U} is less than the payoff of 5 for {I,U}). We assume that the military is facing Knightian uncertainty, that is, the military does not know the probabilities of E or U occurring. If the PMC is not subcontracted, their payoff is 0 regardless of what occurs; if they are subcontracted, they are able to demand additional profit, as explained above, due to problems arising from a failure to assign residual rights. We assume switching costs are high so that the subcontractor has bargaining power. Should the military subcontract the job? It would depend on what it believes is a reasonable chance of U occurring. The more incomplete the contract is to contingencies, the more likely U is to occur. Note that the relative payoffs in the above game may change depending on the arising situation as well: an example of this is the contracting of a reputable PMC that excels at some specialized task. In this case, it could be possible that the payoff of the government under {S,U} could exceed that of {I,U}. Here, playing S would be a strictly dominant strategy, and the government should subcontract the task regardless.

Choosing Tasks to Outsource

If certain tasks must be undertaken in the light of defence policy, the next level of decision is regarding insourcing versus outsourcing for the completion of the tasks. Despite significant economic problems in contracting and related areas such as monitoring and the generation of incentives and the possible lack of competitive markets (see Gosselin (2012) for a detailed analysis of outsourcing in air force), there is much merit in contracting out some aspects of military operations. The operational tasks with the highest potential gains from privatization are those where:

- monitoring is possible and practical;
- performance of the contracted PMC could be compared to some baseline;
- relatively complete contracts can be written.

The inefficiencies associated with PMCs, such as contractual hazards, imperfect monitoring and probity should not be considered in isolation. Despite these problems, outsourcing may still be more efficient than insourcing for a given task.

Navy outsourcing examples

This section includes examples of outsourcing in navies of Australia, the U.K. and the U.S. The examples are specifically drawn from operational areas of naval defence rather than the rather commonly encountered outsourcing in procurement and logistics. The Australian Defence Maritime Services (DMS), the French V.Navy and the British Bristow are examples of PMCs operating at the periphery of naval operations while offering the standard maintenance and logistics services.

DMS technicians maintain and service such mission systems as Rafael Toplite (observation and reconnaissance), Typhoon (mission data exchange) and Mini-Typhoon (remote controlled gun system). It delivers various port and vessel support services under a long-term contract, such as their full support of Armidale class offshore patrol boats. DMS' wide scope operations extend to wherever Australian navy vessels operate. The 65 or so DMS vessels with crews can be contracted to perform such missions as security patrols, marine security barrier maintenance, maritime events support, crewing and vessel delivery, and maritime recovery services. This latter service includes torpedo recovery.

DMS participates in Talisman Sabre, the combined biennial exercise for Australian and US forces in North Queensland. In 2010, the DMS vessel Seahorse Horizon provided mine warfare platform and diving support, recovered "exercise mines from HMAS Waterhen and laid them in the first phase exercise area, ... Meanwhile, USS George Washington, a nuclear-powered aircraft carrier, was anchored off Fremantle before sailing to Darwin to provide the runway for fighter jets on bombing exercises. DMS provided a picket boat to patrol the ship's perimeter while TRV Taylor, operated by RAN crew, monitored the carrier's nuclear emissions. DMS craft supported US Navy supply ships, USNS Alan Shepard and USNS Pecos, which supplied and fuelled by the carrier's crew of 5,000.

As the aircraft carrier left Fremantle, USS Topeka arrived, assisted by DMS tugs, Chuditch and Parma. The Los Angeles Class nuclear powered fast attack submarine and HMAS Farncomb engaged each other in torpedo firing practice in the underwater tracking range, about 90 kilometres off the Western Australian coast. Seahorse Standard was deployed to recover the practice torpedoes." (DMS (2010))

At the "Swiss knife" (or Jack of all trades) end of the spectrum lie the two multi-faceted training facility vessels, VN Rebel operated in the Mediterranean coast of France and VN Partisan operated in the Atlantic coast, owned and manned by the French contractor V.Navy. Capable of training marine commandoes as well as able to land a helicopter and hosting a full mission package, the two vessels serve as sophisticated training facilities for the French forces. Moreover, they can be used in a variety of roles such as targets, boats in distress, mine or torpedo recovery vehicles, towing, landing platform and rescue. For the French forces, this outsourced capability induces economies on expensive regular combat units (V.Navy (2012)).

We will now turn to examples in Australia and Britain where coastal protection and Search & Rescue operations' aerial components have been outsourced. Though not pertaining directly to their navies, those aerial components interface and have to be coordinated with their navies. Australian coastline is

especially vulnerable in the north whereas its southern coasts do not face neighbours. In the late 1980s, the protection of the entire Australian coastline was tasked to the Australian Customs and the mandated division was later named Coastwatch. The Coastwatch Program, initiated in 1995 in response to the need to coordinate naval security introduced partial outsourcing. "Unlike, say, the US Coast Guard, Australia has semi-privatized the coastal patrol function, placing contractors under the Customs service. Once intruders are detected, these contractors can then call on pre-arranged support from civil authorities and/or the Royal Australian Navy and Air Force." The Royal Australian Air Force (RAAF) continues with the mandate to fly sovereignty sorties over Australian coastal waters and coordinates with *Coastwatch* if and when necessary. In addition, a new National Surveillance Centre (NSC) was set up to analyze all incoming information, integrate that information, and make plans for the necessary response. NSC is tasked to tackle this major coordination issue (DeMille (2002)). Coastwatch was re-competed in 2006. Cobham's subsidiary Surveillance Australia Pty Ltd retained the contract under the name Project Sentinel. Under Project Sentinel contract, "Cobham companies will provide, operate and maintain an updated fleet of Dash-8 aircraft through to the year 2021, starting in January 2008. The new services, based on Bombardier Dash-8 aircraft, will double the fleet to provide all-weather, day and night electronic surveillance of Australia's maritime Exclusive Economic Zone." (DeMille (2002)) Thus, coastal aerial surveillance is performed by contractors that interface and coordinate with RAAF (Collins (2013), Defense Industry Daily (2012)).

A different outsourcing structure was recently introduced into the British Search & Rescue (SAR) services. After an open competition, Bristow Helicopters has been awarded the aerial component of the UK SAR contract for Northern Scotland. Bristow's SAR fleet will comprise twenty-two specially commissioned helicopters – specifically the Sikorsky S-92 and the AgustaWestland AW189 replacing the Sea Kings that have previously been used. The company also delivers SAR services "in the Netherlands, Norway, Trinidad, Russia, Brazil, Australia, Canada, Cyprus and the Dutch Antilles as well as providing SAR training in Trinidad and Algeria." (Bristow (2013))

Of the twelve British SAR bases, Bristow will operate the Northern Scottish bases Sumburgh and Stornoway. Thus, unlike the Australian outsourcing of full coastal surveillance, the British outsourcing is localized (Defense Industry Daily (2013b)), perhaps with an eye to the future with full outsourcing.

Analyzed in Gosselin (2012), outsourcing of the aerial component of SAR in Canada is supported by Collins (2013) and DeMille (2010). As successful policy implementations depend on whether structural and motivational conditions exist, one must consider the Canadian environment. The international trend in SAR outsourcing seems to have been strongest for rotary-wing SAR as shown by Bristow's business success (Bristow (2013)). Ironically, Canadian Helicopters International is a major provider of privatized SAR services abroad. But, as in Australia's Coastwatch but there are also examples of private fixed wing services suggesting alternatives to DND's FWSAR project.

The federal Department of Fisheries and Oceans (DFO) has long outsourced surveillance patrols to Provincial Airlines (PAL). Thus a clear trailblazer exists in Canada. If SAR is recognized as a civilian rather than a core military mission, it can be assigned to a civilian agency, such as Transport Canada (TC), in the first place. Existing TC aircraft "operate from civilian airports and are flown by contracted pilots, a situation not unlike that of Australia's Coastwatch. Available CF aircraft still participate in aerial SAR when needed; it just wouldn't invest major resources in SAR squadrons or projects like FWSAR." (CASR (2001)) Thus, a structural transition would require modest steps. The motivational conditions would of course depend on parties' objectives and outsourcing contract properties.

In the United States Navy, outsourcing has been modest and restricted to civilian mariners (CIVMARs) are not are not subject to the military rules and regulations as are Navy personnel. First, the Navy's March 2011 and April 2009 concepts of operations for the Joint High Speed Vessel and Sea-Based X-Band Radar Vessel respectively, include a description of how these ships will be manned by civilian mariners on a rotational basis." (GAO (2008)) The second case is about USS Frank Cable, a submarine support and maintenance vessel. As part of its mission, Frank Cable is to transfer torpedoes, tomahawk missiles, small arms and ammunitions to submarines and surface ships moored alongside at anchorage. The ship's CIVMAR crew must be specially trained and qualified to support these evolutions. (Gcaptain (2011)) The ship also serves as a training facility, somewhat similar to the French V.Navy's two ships considered above. Its sailors train Military Sealift Command's (MSC) CIVMARs on weapons handling and transfer procedures on board the ship. The third case is USS Ponce, which started its commission in 1971 as an "amphibious transport dock ship" and was scheduled to be decommissioned in 2011 but, instead, after being reconfigured as an "afloat forward staging base" re-entered service in 2012, under MSC as USS Frank Cable, to support minesweeping MH-53E Sea Dragon helicopters as part of US Atlantic Fleet. Just like in Frank Cable, a hybrid crew of active-duty Navy officers and sailors, and CIVMARs employed by Military Sealift Command (MSC) operate the ship (Cavas (2012)).

Outsourcing yields a curious tradeoff between the short and long terms. Civilians working as contractors for defence forces earn significantly more than regular soldiers, in fact "about two to ten times more than they would make in their home militaries." (Singer (2005)) For instance, at the right-hand tail of the distribution, a former U.S. Green Beret or member of the Special Forces working in Iraq earned about \$30,000 per month. These high earnings concern military commanders because contractors' recruits come from within the armed services' most highly skilled personnel (Pan (2004)). However, this is a short-term phenomenon as the same high pay will induce young recruits to join armed forces in the long term and acquire the skills necessary expecting to switch to contractors eventually. From the armed forces' perspective, skilled personnel loss is certainly a concern but outsourcing may still reduce costs by reducing the stock of regulars.

As mentioned above in Section 1.5, reservists constitute crucial surge capability. Similarly, PMCs can be thought of similarly in terms of benefits they provide, notwithstanding legal aspects also discussed above in Subsection entitled Legal Considerations but overall lower costs than regular forces as PMCs have to be employed only during those operations for which they are contracted, thus transforming fixed costs and future pension liabilities into simple variable costs.

The military services industry in the USA

The previous section had only one example of outsourcing in the United States. Yet, the current fiscal realities may well force further outsourcing alongside cuts to armed forces. It is thus fitting to take stock of the outsourcing potential in the U.S where the growth in the military services industry has been most obvious. In 2010, the US Department of Defense's annual expenditure on services (including non-military services) accounted for half of the \$400 billion it spent on procurement. Furthermore, the current trend for US arms-producing companies to shift into military services is likely to continue, especially with sequestration as well as entrenched overall budget and debt issues.

On the one hand, the shift is part of strategies to maintain sales in anticipation of cuts in armaments programmes. On the other hand, companies are moving into military services:

(a) to take advantage of general government cost-savings efforts;

(b) to protect themselves from the reduction in projects that had become expected to rotate among the prime contractors as a means of maintaining the financial health of the arms industry (known as Kurth's 'follow-on' imperative); and

(c) to take advantage of governments' general willingness to decrease the number of new programmes and extend the in-service time of existing platforms.

For instance, even before the global financial and economic crisis set in, the US military was planning to move to a commercial maintenance, repair and overhaul (MRO) model for air systems. This 'through-life' approach aims to reduce the costs of aircraft acquisition as well as support for equipment already in service.

There is a general assumption that private companies provide services more cheaply and efficiently than government agencies. This suggests that outsourcing of military services - both through direct government contracts with 'prime contractors' and through third-party contracts from prime contractors - is likely to continue in the short-to-medium term. In addition to potential impacts on costs, these companies are having an impact on security discourse and practice, for example in how equipment maintenance and systems support are provided, what kinds of capabilities countries have in conflict settings and how countries' armed forces are structured and trained (SIPRI (2013)).

While the main drivers behind traditional outsourcing are typically focused on the economic motivators, outsourcing military services can involve a mix of economic, political and military necessities. The common credo of outsourcing is that non-core functions can be outsourced, while core functions — those that define the *raison d'être* of the organization — should never be outsourced. The discussion in the previous section on Canadian SAR sheds light on this question.

In some areas the difficulty is that sometimes the outsourced services are not only important, but are often crucial to the security, and sometimes even survival, of the military authority. Within that context, the contract that defines the relationship between the contractor and the military authority cannot be limited to matters of economic efficiencies, where disputes ultimately get settled between the parties or, less frequently, before the courts. Such contracts must also take into account the different interests, including the loyalties that may exist between the contractor and the military authority in the specific circumstances. This is the probity issue discussed briefly above in the "Market imperfections" paragraph of Section 2.4.

The probity issue surfaces less when outsourcing is farther from the core operational tasks. When logistics services are being outsourced, if the supply to the military authority in operational conditions arrives late or not at all, the military authority will have no problem determining that the contractor is in breach of its obligations under the contract. The result is that the appropriate consequences to the contractor, as stipulated in the contract, will be applied. But what if what's involved is the physical protection of diplomats or a military authority's business sites? Here it may be much more difficult for the military authority to assess compliance with the requirements of the contract. It is extremely important then to consider, before the contract is signed, the military authority's performance expectations, so that non-performance can adequately be assessed, with the appropriate consequences clearly articulated in the contract (Chamberland (2011)). In fact, this ongoing discussion is about the tradeoff between operational effectiveness and cost savings through outsourcing.

Conclusions and extensions

The current study offers an array of examples of cost-saving substitutions from various navies. Since most opportunities for cost-saving substitutions are either technologically driven or derive from new approaches, the study is bound to offer a partial survey. However, a seemingly exhaustive classification of input substitutions has emerged from the examples included.

Arguably the most significant substitution is the one afforded by modularity because navy ships are multi-capability platforms and as such modularity unlocks the traditional vintage capital or equipment trap of system design technologies. Speeding technological innovations can be incorporated into the multi-capability platform as they become available for various functions the platform can perform rather than having to wait for a new class of ships or the mid-life refit of the current class. Moreover, modularity also severs the jointness of various capabilities onboard the platform as a given module, be it a mission or navigation system, can be upgraded relatively independently of others.

Relevant elements of economic analysis were briefly reviewed in Section 1. Understandably, cost minimization lies at the heart of production theory. In the case of a navy, or any other force element for that matter, comparing various input combinations in order to generate capabilities ranging from a platform to a fleet is the essential application. For example, a simple but an important setting is the case of crew rotation. In order to generate warships as part of a fleet, the navy has a choice between assigning a single permanent crew to a ship or a larger set of crews to a smaller set of ships as, for example, in the U.S. Navy's Blue and Gold crews assigned to Trident submarines. The main issue arising in this context is whether longer periods of a vessel's deployment enabled by crew rotation generates lower overall costs than generating the same capability with single crew but more numerous vessels. Important cost considerations are the higher maintenance costs of a vessel due to sustained deployment and incremental costs of managing further crews trading off against having a smaller number of vessels with perhaps higher procurement costs due to shorter production runs. Back of the envelope calculations wouldn't suffice as additional costs must be considered. Further consideration of costs is required since navy platforms last several decades and risks of strategic obsolescence must be internalized. Though alleviated by modularity in design, such risks augment costs due to other relevant risks. Moreover, if the vessel and fleet in question are supported by outsourced capabilities, as in the case of U.S. Military Sealift Commander of U.K.'s Search & Rescue, then significant transaction costs associated with outsourcing contracts will arise and have to be incorporated into decision-making.

This study spans navy input substitutions amongst the existing range of inputs within the navy and across governmental organizations, those enabled by product design and, lastly, by outsourcing identified as organizational substitution. Outsourcing may not only change input proportions from insourced to outsourced component of the organization but also changes the intrinsic nature of costs. Whereas in-house production generates fixed costs, outsourcing transforms most fixed costs into variable costs through contracting. As such, outsourcing can easily be the most complex of all substitutions considered.

Apart from well-researched issues and ongoing debate on certain substitutions, new questions have emerged from this study. First, as the above paragraph demonstrates, interactions between substitutions induce trade-offs in terms of cost savings. These trade-offs necessarily involve intertemporal considerations, not only in terms of a simple arithmetic of costs but also issues related to operational capabilities. For instance, outsourcing an operational capability may generate a force hollowing out with the consequence that a lessening of competition on the supply side may jeopardize operational effectiveness in terms of timely surges. Second, there may exist a scale dependency in terms

of feasible substitutions or, in other words, larger navies may have a wider spectrum of possibilities. For example, given the strategic environment, a larger navy may make use of a larger number of smaller or simpler vessels instead of bigger or more complicated ones but economizing nevertheless from procurement cost savings due to larger production runs. Yet another example is the organizational substitution with U.S. Navy's surface warfare officer retention auctions (Coughlan et al. (2011)) where the existence of a large group of officers and a large number of billets make it possible to run combinatorial auctions to match officers with postings while saving overall costs.

References

- Abbott, J.W. (2006), "The History of Modular Payload Ships: 1975 - 2005", Naval Postgraduate School, <http://www.nps.edu/Academics/Institutes/Meyer/docs/April%2027%202006%20History%20of%20Modular%20Payload%20ships.pdf>, accessed 05 July 2013
- Ailes, J. (2013), "LCS Mission Modules", US Navy, <http://www.navsea.navy.mil/Media/SAS2013/3.%20LCS%20Mission%20packages.pdf>, accessed 05 July 2013
- Atherton, K.D. (2013), "[U.S. Navy Spends \\$37 Billion On A Ship That Barely Works](http://www.popsi.com/technology/article/2013-05/us-navy-spends-37-billion-ship-that-barely-works)", <http://www.popsi.com/technology/article/2013-05/us-navy-spends-37-billion-ship-that-barely-works>, accessed 05 July 2013
- Berthiaume, L. (2012), "Canadian military building network of international outposts in push for more overseas missions", *National Post*, 19 July 2012
- Bosworth, M.L. et al. (1994), "[Multimission Ship Design for an Alternative Fleet Concept](#)", *Naval Engineers Journal* 103(3), 91-106
- Bristow (2013), "About us", <http://bristowsar.com/>, accessed 04 August 2013
- Bucci, V., A. Marinò & I. Juricic [2013], "Integrated ship design: Automated Generation of Production Deliverables with New Generation Shipbuilding CAD Systems", paper presented at 12th International Conference on Computer and IT Applications in the Maritime Industries, Cortona, 15-17 April 2013 http://www.ssi.tu-harburg.de/doc/webseiten_dokumente/compit/dokumente/Paper2013_Bucci.pdf, accessed 15 June 2013
- Cameron, L. (2006), "Private military companies: their status under international humanitarian law and its impact on their regulation", *International Review of the Red Cross* 88(863), 573-598
- Canadian Forces (2013), "Operational Support Hubs", <http://www.forces.gc.ca/en/operations-support/os-hubs.page>, accessed 04 August 2013
- Caprace, J.-D. (2010), "Cost effectiveness and complexity assessment in ship design within a concurrent engineering and 'Design for X' framework", PhD Thesis, Université de Liège
- Carter, A.B. & J.C. Bulkeley (2007), "America's Strategic Response to China's Military Modernization", *Harvard Asia Pacific Review* 9(1), 50-52

CASR (2001), "A Modest Proposal - SAR Aircraft: Rethinking Search and Rescue", Canadian American Strategic Review (CASR), <http://www.casr.ca/mp-fwsar-alternative.htm>, accessed 31 July 2013

Cavas, C.P. (2012), "New Floating Base Ships Coming for U.S. Navy", <http://www.defensenews.com/article/20120127/DEFREG02/301270010/New-Floating-Base-Ships-Coming-U-S-Navy>, accessed 26 July 2013

CBO (Congressional Budget Office) (2007), "Crew Rotation in the Navy: The Long-Term Effect on Forward Presence", Congressional Budget Office, Congress of the United States

CBO (Congressional Budget Office) (2006), "Options for the Navy's future fleet", Congressional Budget Office, Congress of the United States

Chamberland, D. (2011), "Contractors on the battlefield: Outsourcing of military services", National Defense Magazine, <http://www.nationaldefensemagazine.org/archive/2011/March/Pages/ContractorsontheBattlefieldOutsourcingofMilitaryServices.aspx>, accessed 29 July 2013

Christopoulou, K., C. Makatsoris & A. Shortland (2013), "War and Famine, Peace and Light?", Working Paper presented at the 7th Defence & Security Economics Workshop, DRDC and RMCC, Ottawa

Chuter, A. (2013), "U.K. Privatized Procurement Moves Forward", <http://www.defensenews.com/article/20130330/DEFREG01/303300007/U-K-Privatized-Procurement-Moves-Forward>, accessed 26 July 2013

Coase, R. H. (1937), "The Nature of the Firm", *Economica* 4(16), 386-405

Collette, M. (2011), "Modularity and ship retirement: Hull Structures as a System: Supporting Lifecycle Analysis", *Naval Engineers Journal* 123(3), 45-55

Collins, M. (2013), "Contracting our Rotary SAR? Plus Australia's Coastwatch", <http://www.cdfai.org/the3dsblog/?p=1724>, accessed 04 August 2013

Coughlan, P.J., W.R. Gates & B.M. Zimmerman (2011), "The Combinatorial Retention Auction Mechanism (CRAM): Integrating Monetary and Non-Monetary Reenlistment Incentives", Naval Postgraduate School, NPS-GSBPP-11-013, <http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA556107>

Covert, D. (2012), "Modularity: Naval ship modernization with reduced cost and increased flexibility", <http://www.10thsymposium.com/presentations/Wed%20pm%20A/1630%20Covert-%20MIW%20Tech%20Symposium%20Brief.pdf>, Mine Warfare Technology Symposium, 09 May 2012, accessed 04 July 2013

Damen Schelde (2013), "SIGMA series", http://www.damennaval.com/nl/main-stream-activities_combatants-and-opv%2527s_sigma-series.htm, accessed 05 July 2013

Defense Industry Daily (2013a), "LCS: The USA's Littoral Combat Ships", <http://www.defenseindustrydaily.com/the-usas-new-littoral-combat-ships-updated-01343/>, accessed 26 June 2013

Defense Industry Daily (2013b), "Britain's Next Search-and-Rescue Helicopters: Civilian Contractors", <http://www.defenseindustrydaily.com/british-searchandrescue-a-billion-pound-partnership-02271/>, accessed 31 July 2013

Defense Industry Daily (2012), "Australia's Coastwatch: A Public-Private Model for Coast Guards and CBP", <http://www.defenseindustrydaily.com/cobham-catches-a-1b-australian-coastwatch-contract-01695/>

Defence Management (2013), "Privatising defence procurement still 'preferred' over civil service", http://www.defencemanagement.com/news_story.asp?id=22120, accessed 22 July 2013

DeMille, D. (2010), "Air Force must focus on core missions — Time to calve off SAR Local Civilian Agencies may prove to be more Cost-Effective", CASR, <http://www.casr.ca/ft-sar-civilian-1.htm>

DeMille, D. (2002), "Australia's 'Coastwatch' - What Can Canada Learn?", CASR, <http://www.casr.ca/ft-coastwatch-australia.htm>

DMS (2010), "DMS Year in Review", http://www.dmsmaritime.com/docs/DMS%20YIR_web.pdf

Doerry, N.H. (2012), "Institutionalizing Modular Adaptable Ship Technologies", Paper No. SNAME-047-2012, <http://www.sname.org/Go.aspx?NavigationKey=7f209ba5-e131-4648-b324-a2f43404edeb>, accessed 15 June 2013

Duffield, M. (1994), "The political economy of internal war: asset transfer, complex emergencies and international aid", in J. Macrae & A. Zwi (Eds.), War and hunger: rethinking international responses to complex emergencies, London: Zed Books

European Commission (2000), "Adapting robots to European shipbuilding" <http://ec.europa.eu/research/growth/gcc/projects/robmar.html>, accessed 26 June 2013

Famme, J. (1994), "Automation: The Use of Automated Ship Control Systems Technology to Reduce the Cost of Ships and Submarines", ASNE Intelligent Ship Symposium 1994, <http://www.businessdevelopmentusa.com/references/Microsoft%20Word%20-%20ShipSystemsAutomationTechnologyReprintMay2004MASTER2.doc.pdf>, accessed 19 July 2013

Fernholtz, T. (2013), "The outsourcing debate behind the Dreamliner debacle—and the memo Boeing execs should have read", <http://finance.yahoo.com/news/outsourcing-debate-behind-dreamliner-debacle-110057532.html>, accessed 31 July 2013

Meagan Fitzpatrick, M. (2012), "Canadian Forces to open support hub in Germany", CBC News, <http://www.cbc.ca/news/world/story/2012/02/14/pol-mackay-germany.html>

Franck, R. & F. Mélése (2008), "Defence Acquisition: New Insights from Transaction Cost Economics", Defence and Security Analysis 24(2), 107-128

Gale, D. & L.S. Shapley (1962), "College admissions and the stability of marriage", American Mathematical Monthly 69(1), 9-15

GAO (2013), "Defense Contracting: Actions Needed to Increase Competition", US Government Accountability Office, GAO-13-325

GAO (2010), "Military Readiness: Navy Needs to Reassess Its Metrics and Assumptions for Ship Crewing Requirements and Training", Report to Congressional Committees

GAO (2008), "Ship Rotational Crewing Initiatives Would Benefit from Top-Level Leadership, Navy-wide Guidance, Comprehensive Analysis, and Improved Lessons-Learned Sharing", GAO-08-418

GAO (2005), "Defense Acquisitions: Improved Management Practices Could Help Minimize Cost Growth in Navy Shipbuilding Programs", Report to the Chairman, Subcommittee on Defense, Committee on Appropriations, House of Representatives, GAO-05-183

GAO (2005), "Interagency Contracting: Problems with DOD's and Interior's Orders to Support Military Operations", US Government Accountability Office, GAO-05-201

GAO (2003), "Military Personnel Navy - Action Needed to Optimize Ship Crew Size and Reduce Total Ownership Costs", GAO-03-520

Gcaptain (2011), "Military Sealift Command civilian mariners trained on weapons handling and transfer procedures", <http://gcaptain.com/military-sealift-command-civilian/>, accessed 31 July 2013

Gelling, J.L. & L. Goossens (2008), "Damen's philosophy: Standardized patrol boats and naval vessels", Engineering the Total Ship (ETS) 2008 Proceedings, American Society of Naval Engineers (ASNE), <https://www.navalengineers.org/publications/symposiaproceedings/Proceedings2008/Pages/ETS2008Proceedings.aspx>, accessed 05 July 2013

Gifford, S. (2008), "The Armed Private Military Company and the Canadian Forces: The Next Step in Contractor Support?", *Canadian Army Journal* 11(1), 78-97

Gillet, M. (2013) "The Anatomy of an International Crime: Aggression at the International Criminal Court", *International Criminal Law Review*, <http://ssrn.com/abstract=2209687> or <http://dx.doi.org/10.2139/ssrn.2209687>

Girard, S. et al. (2008), "Canadian Forces overseas supply network: Strategic need and design methodology", CIRRELT-2008-34, <https://www.cirrelt.ca/DocumentsTravail/CIRRELT-2008-34.pdf>

Global Security (2013), "Sea Swap / Crew Rotation", <http://www.globalsecurity.org/military/ops/sea-swap.htm>, accessed 14 July 2013

Golding, H.L.W. & G.E. Cox (2003), "Design and Implementation of AIP", The CNA Corporation, CAB D0007827.A2/Final, <http://www.cna.org/sites/default/files/research/d0007827.a2.pdf>

Gosselin, D.G. (2012), "Outsourcing Operational Effectiveness? A Case Study of the Royal Canadian Air Force", MPA Thesis, Royal Military College of Canada

Grimes, A. & J. Rolfe (2002), "Optimal defence structure for a small country", *Defence and Peace Economics* 13(4), 271-286

Guttman, D. (2004), "Governance by Contract: Constitutional Visions; Time for Reflection and Choice", *Public Contract Law Journal* 33(2), 321-360

Hagan, J. et al. (2011), "Human Systems Integration and Crew Design Process Development in the Zumwalt Destroyer Program", Special Report 306, Paper prepared for the Committee on Naval

- Engineering in the 21st Century Transportation Research Board,
<http://onlinepubs.trb.org/onlinepubs/nec/6810Hagan.pdf>, accessed 13 July 2013
- Halonen-Akatwijuka, M. & O.D. Hart (2013), "More is less: Why parties may deliberately write incomplete contracts", NBER WP 19001, <http://www.nber.org/papers/w19001>
- Hartley, K. (2004), "The Economics of Military Outsourcing", *Defence Studies* 4(2), 199-206
- Hildebrandt, G.G. (1999), "The Military Production Function", *Defence and Peace Economics* 10, 247-272
- Hoffman (2006), "The fleet we need: A look at alternative — and affordable — futures for the U.S. Navy", <http://www.armedforcesjournal.com/2006/08/1891359/>, accessed 11 July 2013
- Jehle, G.A. & P.J. Reny (2011), *Advanced Microeconomic Theory*, 3rd Edition, Prentice Hall
- Johnson, D.C. et al. (2009), "Managing Change on Complex Programs: VIRGINIA Class Cost Reduction," *Naval Engineers Journal* 121(4), 79-94
- Johnson, S.E. & A.K. Cebrowski (2005), "Alternative Fleet Architecture Design", Center for Technology and National Security Policy, National Defense University,
<http://www.ndu.edu/CTNSP/docUploaded/DTP%2019%20Alternative%20Fleet%20Architecture%20Design.pdf>, accessed 31 May 2013
- Koenig, P. (2009), "Real Options in Ship and Force Structure Analysis: A Research Agenda," *Naval Engineers Journal* 121(4), pp. 95-105
- Leland, H.E. (1968), "Savings and Uncertainty: The Precautionary Demand for Saving", *Quarterly Journal of Economics* 82(3), 465-473
- Lyridis, D.V. (2005), "Introduction to an Innovative Crew Composition Approach Based on Safety/Operational and Financial Requirements", *WMU Journal of Maritime Affairs* 4(1), 35-57
- Mas-Colell, A., M.D. Winston & J.R. Green (1995), *Microeconomic Theory*, Oxford University Press
- Milliard, T. S. (2003), "Overcoming Post-Colonial Myopia: A Call to Recognize and Regular Private Military Companies", *Military Law Review* 176(June), 1-95.
- Minow, M. (2005), "Outsourcing Power: How Privatizing Military Efforts Challenges Accountability, Professionalism, and Democracy", *Boston College Law Review* 46(5), 989-1026
- Mobley, B.W (2004), "Outsourcing post-conflict operations: Designing a system for contract management and oversight", *Journal of Public and International Affairs* 15, 21-37
- NAO (National Audit Office) (2010), "Delivering multi-role tanker aircraft capability", HC 433, U.K. National Audit Office, <http://www.nao.org.uk/wp-content/uploads/2010/03/0910433.pdf>
- Naval Technology (2013), "U212 / U214 Submarines, Germany", http://www.naval-technology.com/projects/type_212/, accessed 04 July 2013
- Naval Technology (2012), "Buques de Acción Marítima (BAM) Class Patrol Vessels, Spain", <http://www.naval-technology.com/projects/maritimeactionship/>, accessed 05 July 2013

- O'Brien, K.A. (2002) "Private Military Companies: Options for Regulation", RAND Europe, http://www.ssrnetwork.net/uploaded_files/3541.pdf
- O'Hanlon, M. (2013), "How to cut the Pentagon budget better than sequestration does", Washington Post, 12 May 2013, http://articles.washingtonpost.com/2013-05-12/opinions/39210831_1_automatic-spending-cuts-further-defense-cuts-sequestration
- O'Rourke, R. (2013), "Navy Littoral Combat Ship (LCS) Program: Background and Issues for Congress", CRS Report for Congress, Order Code RL33741
- O'Rourke, R. (2007), "Navy Force Structure: Alternative Force Structure Studies of 2005 — Background for Congress", Congressional Research Service RL 33955
- O'Rourke, R. (2005), "Navy Ship Acquisition: Options for Lower-Cost Ship Designs - Issues for Congress", CRS Report for Congress, Order Code RL32914
- Pan, E. (2004), "Iraq: Military Outsourcing", Council on Foreign Relations, <http://www.cfr.org/security-contractors/iraq-military-outsourcing/p7667>
- Pernin, C.G. et al. (2011), "Efficiencies from Applying a Rotational Equipping Strategy", RAND, http://www.rand.org/content/dam/rand/pubs/monographs/2011/RAND_MG1092.pdf
- Powers, R. (2013), "Navy Enlisted Rating (Job) Descriptions and Qualification Factors Damage Controlman (DC)", *About.com US Military*, <http://usmilitary.about.com/od/enlistedjob1/a/dc.htm>, accessed 08 July 2013
- Rohlf, C. (2011), "How Much Did the United States Government Value its Troops' Lives in World War II? Evidence from Dollar-Fatality Tradeoffs in Ground Warfare", Center for Policy Research, Syracuse University, https://carohlf.mysite.syr.edu/tanks_020511.pdf
- Rohlf, C. (2006), "The Government's Valuation of Military Life-Saving in War: A Cost-Minimization Approach", *American Economic Review* 96(2), 39-44
- Rostek, M.A. (2006), "Developing a Surge Capacity for Canadian Forces", *Defence and Peace Economics* 17(5), 421-434
- Roth, A.E. & M. Sotomayor (1990), *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*, Cambridge University Press
- Rousseau, R. (2012), "Are Private Military Companies Exempted from Geneva Conventions?", *Diplomatic Courier*, 23 November 2012, <http://www.globalpolicy.org/component/content/article/257-pmscs-risks-misconduct/52093-are-private-military-companies-exempted-from-geneva-conventions.html>
- Schooner, S. L. (2005), "Contractor Atrocities at Abu Ghraib: Compromised Accountability in a Streamlined, Outsourced Government", *Stanford Law and Policy Review* 16, 549-572
- Schwartz, M. & J. Swain (2011), "Contractors in Afghanistan and Iraq: Background and Analysis", Congressional Research Service, R40764
- Scofield, T.J. & A.J. Brown (2007), "Manning and automation model for naval ship analysis and optimization", ASNE Human Systems Integration Symposium, Annapolis MD, 2007

- Singer, P. W. (2003), "Peacekeepers, Inc.", Policy Review 119(June)
- Singer, P. W. (2005), "Outsourcing War", Foreign Affairs 84(2), 119-132
- SIPRI (2013), "Private Military Services Industry", Stockholm International Peace Research Institute, <http://www.sipri.org/research/armaments/production/researchissues/pmsc>, accessed 29 July 2013
- Sönmez, T. and T.B. Switzer (2013), "Matching with (Branch-of-Choice) Contracts in the United States Military Academy", Econometrica 81(2), 451–488
- Stoker, L. (2012), "Combat ships do battle: LCS versus Type 26", <http://www.naval-technology.com/features/featurecombat-ships-battle-lcs-type-26>, accessed 05 July 2013
- Strategy Page (2011), "Why The U.S. Navy Is Falling Apart", www.strategypage.com/htm/htsurf/articles/20110715.aspx, accessed 19 July 2013
- Sweeney, C. (2010), "The World's 18 Strangest Military Bases", Popular Mechanics, <http://www.popularmechanics.com/technology/engineering/architecture/strangest-military-bases-gallery>, accessed 04 August 2013
- Tavik, G. et al. (2010), "Integrated Topside (InTop) Joint Navy-Industry Open Architecture Study" NRL/FR/5008--10-10,198, Sept 20, 2010
- U.K. Green Paper (2002), "Private Military Companies: Options for Regulation 2001-02", HC577 Foreign and Commonwealth Office, House of Commons, <http://www.globalsecurity.org/military/library/report/2002/mercenaries.pdf>
- U.S. Army (2013), "Career Satisfaction Program", <http://www.career-satisfaction.army.mil/index.html>, accessed 27 July 2013
- V. Navy (2012), "VN Partisan: Un couteau suisse pour l'entraînement des marins", <http://www.meretmarine.com/fr/content/vn-partisan-un-couteau-suisse-pour-lentrainement-des-marins>
- Vickrey, W. (1961), "Counterspeculation, Auctions, and Competitive Sealed Tenders", Journal of Finance 16(1), 8-37
- Viscusi, W.K. (2013), "Estimating the Value of a Statistical Life Using Census of Fatal Occupational Injuries (CFOI) Data", Vanderbilt University Law School, Law and Economics Working Paper Number 13-17, <http://ssrn.com/abstract=2291727>
- Volkert, R., C. Jackson & C. Whitfield (2010), "Development of Modular Mission Packages Providing Focused Warfighting Capability for the Littoral Combat Ship", Naval Engineers Journal 122(4), 75-92
- Williamson, O.E. (1999), "Public and Private Bureaucracies: A Transaction Cost Economics Perspective", Journal of Law, Economics & Organization 15(1), 306-342
- Williamson, O. E. (1981), "The Economics of Organization: The Transaction Cost Approach", American Journal of Sociology 87, 548-577